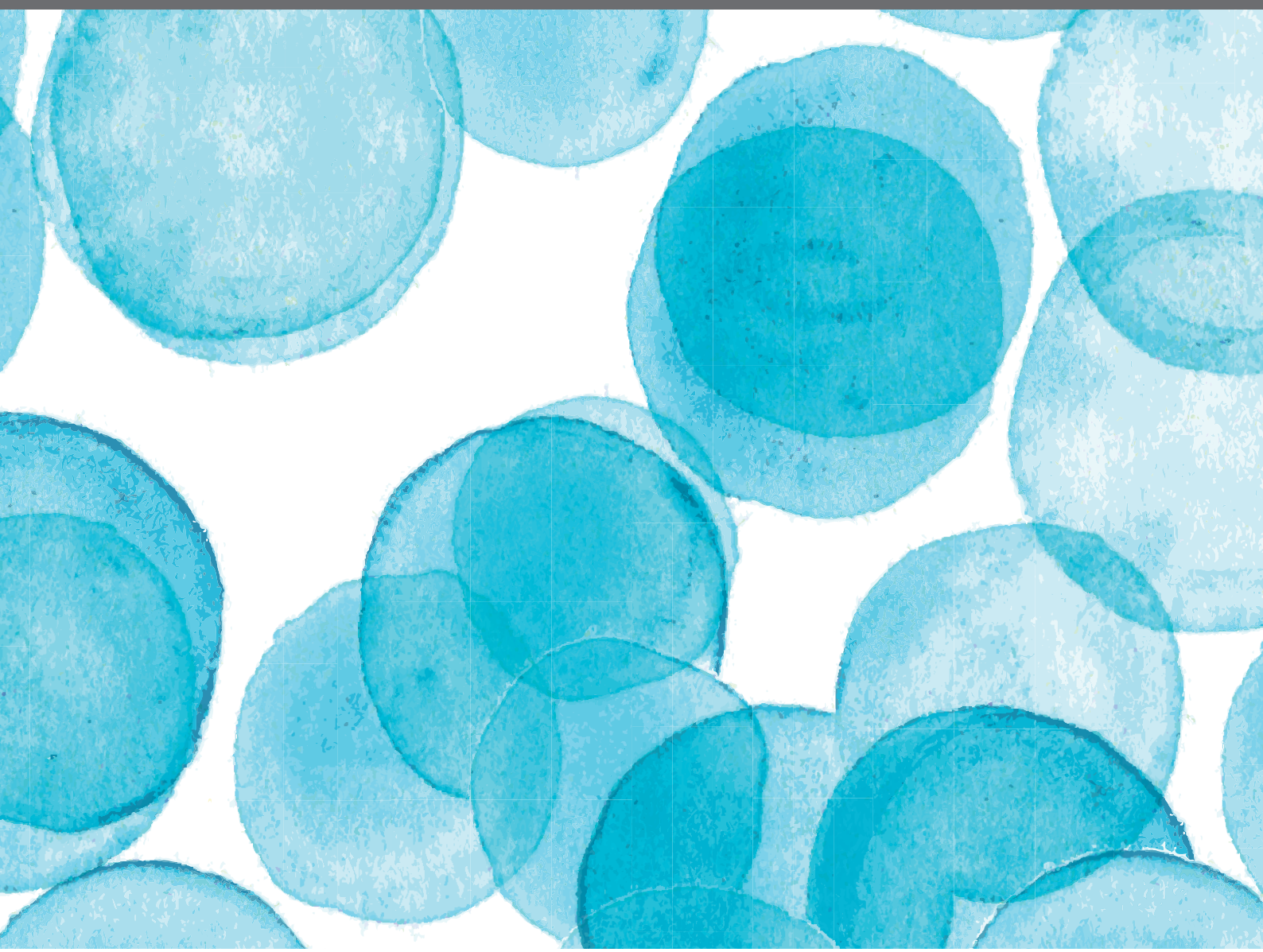
ADVANCES IN MULTI-OMICS STUDY OF FILAMENTOUS PLANT PATHOGENS

EDITED BY: Danyu Shen, Yan Wang, Xiao-Ren Chen, Vaibhav Srivastava and Silvia Laura Toffolatti

PUBLISHED IN: Frontiers in Microbiology



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-83250-034-7
DOI 10.3389/978-2-83250-034-7

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

ADVANCES IN MULTI-OMICS STUDY OF FILAMENTOUS PLANT PATHOGENS

Topic Editors:
Danyu Shen, Nanjing Agricultural University, China
Yan Wang, Nanjing Agricultural University, China
Xiao-Ren Chen, Yangzhou University, China
Vaibhav Srivastava, Royal Institute of Technology, Sweden
Silvia Laura Toffolatti, University of Milan, Italy

Citation: Shen, D., Wang, Y., Chen, X.-R., Srivastava, V., Toffolatti, S. L., eds. (2022). Advances in Multi-omics Study of Filamentous Plant Pathogens. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-034-7

Table of Contents

frontiers | Frontiers in Microbiology

TYPE Editorial
PUBLISHED 12 August 2022
DOI 10.3389/fmicb.2022.998501

OPEN ACCESS

EDITED AND REVIEWED BY
Ludmila Chistoserdova,
University of Washington,
United States

*CORRESPONDENCE
Danyu Shen
shendanyu@njau.edu.cn
Yan Wang
yan.wang@njau.edu.cn
Xiaoren Chen
xrchen@yzu.edu.cn
Vaibhav Srivastava
vasri@kth.se
Silvia Laura Toffolatti
silvia.toffolatti@unimi.it

SPECIALTY SECTION
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

RECEIVED 20 July 2022
ACCEPTED 01 August 2022
PUBLISHED 12 August 2022

CITATION
Shen D, Wang Y, Chen X, Srivastava V and Toffolatti SL (2022) Editorial: Advances in multi-omics study of filamentous plant pathogens. *Front. Microbiol.* 13:998501. doi: 10.3389/fmicb.2022.998501

Editorial: Advances in multi-omics study of filamentous plant pathogens

Danyu Shen[1]*, Yan Wang[1]*, Xiaoren Chen[2]*, Vaibhav Srivastava[3]* and Silvia Laura Toffolatti[4]*

[1]Department of Plant Pathology, Nanjing Agricultural University, Nanjing, China, [2]College of Horticulture and Plant Protection, Yangzhou University, Yangzhou, China, [3]Division of Glycoscience, Department of Chemistry, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, AlbaNova University Center, Stockholm, Sweden, [4]Dipartimento di Scienze Agrarie ed Ambientali, Università degli Studi di Milano, Milan, Italy

KEYWORDS

omics, oomycetes, fungi, pathogenesis, plant-pathogen interaction

Editorial on the Research Topic
Advances in multi-omics study of filamentous plant pathogens

Filamentous plant pathogens, such as fungi and oomycetes, cause devastating diseases of crop plants, and result in serious threats to agriculture and natural ecosystems worldwide. Diseases caused by fungal pathogens, such as wheat stem rust (*Puccinia graminis*), rice blast (*Magnaporthe oryzae*), and gray mold (*Botrytis cinerea*), are of immediate concern for global food security. The oomycete *Phytophthora infestans*, the causal agent of the Great Irish Famine in the nineteenth century, remains a recurring threat to potato and tomato production today. *P. sojae* causes soybean root and stem rot, while *P. capsici* infects a large number of agriculturally important vegetables, resulting in severe yield losses. In order to control such plant diseases, the broad spectrum of the infection process and pathogenic mechanisms that filamentous pathogens apply to colonize plants needs to be elucidated. However, there are still many gaps in our current understanding. High-throughput technologies have revolutionized plant pathology research. Omics technologies, such as genomics, transcriptomics, proteomics, metabolomics, epigenomics, and microbiomics, have become essential resources in the plant-pathogen interaction research. Pathogen genomics focuses on the structure, function, and evolution of genomes, while transcriptomics and proteomics provide insights into gene and protein expression. This Research Topic aimed to collect studies of omics data to provide an increased understanding of various aspects related to plant-pathogen biology and plant-pathogen interactions. It comprises nine interesting papers covering a broad range of themes. Here, we grouped these articles into the following five themes.

Genomics

Whole-genome sequencing is a comprehensive approach for analyzing entire genomes, and thus is significantly helpful for biological research. Mandal et al. analyzed 128 *Phytophthora* genomes, and found that the simple sequence repeats of all studied *Phytophthora* species followed distinct isolate-specific patterns. The pathogenesis-related genes were localized in more repeat-rich regions, and RxLR effectors diverged rapidly and did not show any common core group. Nellist et al. performed comparative genomic analysis of 18 *Phytophthora cactorum* isolates, and these isolates showed specialization in strawberry crown and apple. This study also highlighted several RxLR effectors that warrant further investigation as host specificity determinants. Wyrebek et al. evaluated the value of mitogenome to identify *Fusarium graminearum sensu stricto* (*F.g.*). They found that homing endonucleases displayed small indel polymorphism, which facilitated further distinguishing between *F.g.* and related species belonging to *F. graminearum* species complex.

Transcriptomics

Transcriptome sequencing provides data on the expression of almost all transcripts in a certain state, and has been widely used in the study of plant-pathogen interactions. Daly et al. performed a dual-transcriptomic analysis of the interaction between the biocontrol agent *Pythium oligandrum* and ginger pathogen *Pythium myriotylum*, and found that *Py. myriotylum* genes encoding Kazal-type protease inhibitors, cellulases, and elicitin-like proteins were strongly up-regulated, while *Py. oligandrum* genes coding for proteases, cellulases, and peroxidases were highly activated. Tang et al. investigated transcriptome variations of *Verticillium dahliae* in response to two different inorganic nitrogen sources. They found that genes related to oxidoreductase, cell cycle, and some metabolic activities were down-regulated in the ammonium treatment. Meanwhile, differentially expressed genes related to glycerol transport, energy-dependent multidrug efflux pump activity, and L-serine biosynthesis, were involved in the utilization of both nitrate and ammonium.

Proteomics

Post-translational modifications, such as redox modification, are involved in many biological processes and physiological pathways in both eukaryotes and prokaryotes. However, protein redox modification in phytopathogenic fungi is unclear. Zhang et al. surveyed thiol proteome using a mixed sample containing mycelia with or without oxidative stress, invasive hyphae, conidia, and appressoria of the model fungal pathogen *Magnaporthe oryzae,* and identified the thiol-modified proteins by protein domain, functional classification, subcellular localization, metabolic pathways, and protein–protein interaction network analyses. Their findings indicated that redox modification may play key roles in fungal growth, conidium formation, appressorium formation, and invasive growth. This study provides a global insight into the redox proteome of the plant pathogenic fungi and is a valuable resource for future studies on redox modification in fungi.

Metabolomics

Metabolism is the set of biochemical reactions that occur in living organisms. During plant-microbe interactions, extensive changes occur in plant and microbe metabolism involved in many molecular and cellular processes taking place within cells. Genome-scale metabolic models (GEMs) have become an important tool for systematically analyzing metabolic alterations associated with plant-microbe interactions. Due to their integrative nature, by relating genes with reactions and metabolites, GEMs can simulate the system-wide metabolic fluxes of both the plant and pathogen, and help identify genes, reactions, and metabolites essential for understanding the mechanisms underlying plant-microbe interactions. Rodenburg et al. reviewed the current knowledge applying metabolic modeling to dissect microbial pathogenesis and host–pathogen interactions and proposed a workflow for reconstructing high-quality GEMs.

Functional studies

The mevalonate (MVA) pathway is involved in the biosynthesis of isoprenoids and sterols, but the information is limited in the sterol-auxotrophic *Phytophthora* species. Yang et al. firstly identified the whole set of MVA pathway genes in *Phytophthora sojae in silico*. Then, they explored the functions of MVA pathway by treatment with enzyme inhibitor lovastatin, deletion of the geranylgeranyl diphosphate synthase gene *PsBTS1*, and transcriptome analysis. They discovered that the MVA pathway played essential roles in vegetative growth, development, sexual and asexual reproduction and virulence in *P. sojae*. He et al. characterized the functions of the transcription factor *VpxlnR* in *Valsa pyri*. Through creating gene deletion mutants, they discovered that *VpxlnR* was involved in fruiting body differentiation, growth on culture media, response to chemical stress (hydrogen peroxide and salicylic acid) and infection of detached pear leaves and branches at the early infection stages. Furthermore, they identified five target genes that could interact with VpxlnR *in vivo* based on a yeast one-hybrid assay, providing

useful information about the mechanism of pathogenesis of *V. pyri*.

Author contributions

DS, YW, XC, VS, and ST co-edited the Research Topic and wrote, edited, and approved the final version of the editorial.

Funding

DS was funded by the National Natural Science Foundation of China (NSFC; 32070139), and Jiangsu Agricultural Science and Technology Innovation Fund [CX(21)3085]. XC was funded by NSFC (31871907, 31671971) and Jiangsu Agriculture Science and Technology Innovation Fund (JASTIF) (CX(20)3125). VS was funded by European Commission (Horizon 2020 FET-OPEN Grant Agreement No. 828940) and the Swedish Research Council FORMAS (Grant #2019-00912).

Acknowledgments

We thank the Frontiers Editorial Office and reviewers for their assistance in completing this Research Topic.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

frontiers in Microbiology

ORIGINAL RESEARCH
published: 10 March 2021
doi: 10.3389/fmicb.2021.648894

The Redox Proteome of Thiol Proteins in the Rice Blast Fungus *Magnaporthe oryzae*

Xinrong Zhang[1,2†], Zhenhua Zhang[1,3†] and Xiao-Lin Chen[1]*

[1] State Key Laboratory of Agricultural Microbiology, Provincial Key Laboratory of Plant Pathology of Hubei Province, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, [2] State Key Laboratory of Agrobiotechnology, Ministry of Agriculture Key Laboratory for Plant Pathology, China Agricultural University, Beijing, China, [3] Department of Genetics, University Medical Center Groningen, Groningen, Netherlands

Redox modification, a post-translational modification, has been demonstrated to be significant for many physiological pathways and biological processes in both eukaryotes and prokaryotes. However, little is known about the global profile of protein redox modification in fungi. To explore the roles of redox modification in the plant pathogenic fungi, a global thiol proteome survey was performed in the model fungal pathogen *Magnaporthe oryzae*. A total of 3713 redox modification sites from 1899 proteins were identified through a mix sample containing mycelia with or without oxidative stress, conidia, appressoria, and invasive hyphae of *M. oryzae*. The identified thiol-modified proteins were performed with protein domain, subcellular localization, functional classification, metabolic pathways, and protein–protein interaction network analyses, indicating that redox modification is associated with a wide range of biological and cellular functions. These results suggested that redox modification plays important roles in fungal growth, conidium formation, appressorium formation, as well as invasive growth. Interestingly, a large number of pathogenesis-related proteins were redox modification targets, suggesting the significant roles of redox modification in pathogenicity of *M. oryzae*. This work provides a global insight into the redox proteome of the pathogenic fungi, which built a groundwork and valuable resource for future studies of redox modification in fungi.

Keywords: post-translational modification, *Magnaporthe oryzae*, redox proteome, thiol proteins, fungal infection, oxidative stress

OPEN ACCESS

Edited by:
Xiao-Ren Chen,
Yangzhou University, China

Reviewed by:
Min He,
Sichuan Agricultural University, China
Min Guo,
Anhui Agricultural University, China

***Correspondence:**
Xiao-Lin Chen
chenxiaolin@mail.hzau.edu.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *02 January 2021*
Accepted: *28 January 2021*
Published: *10 March 2021*

Citation:
Zhang X, Zhang Z and Chen X-L (2021) The Redox Proteome of Thiol Proteins in the Rice Blast Fungus Magnaporthe oryzae. Front. Microbiol. 12:648894. doi: 10.3389/fmicb.2021.648894

INTRODUCTION

Protein features and functions are modulated not only by sequence of the amino acids but also by post-translational modifications (PTMs), such as acetylation, lipidation, methylation, and oxidation–reduction modification (as known as redox modification). PTMs are ideal approach to maintain the cell homeostasis and response to external or internal stimuli by recasting the function, stability, or activity of translated proteins, which consequently impacts the organism at different levels such as genome and epigenome (Go and Jones, 2013).

In PTMs, the redox modification on cysteine residues is an important source of regulations on protein functions (Go and Jones, 2013; Sievers et al., 2018), which impacts many physiological pathways and biological processes (Zhao and Jensen, 2009; Baraibar and Friguet, 2013; Olsen and Mann, 2013).

The reactive thiol groups attached to cysteine residues are responsible to the redox modifications by shifting from reduction to oxidation or conversely (Winterbourn and Hampton, 2008; Hancock, 2009; Paulsen and Carroll, 2010), where the shifting is either irreversible (e.g., sulfonic acid) or reversible (e.g., disulfide bonds) (Murray and Van Eyk, 2012). In previous studies, it has been revealed that protein structures and functions can be transformed by reversible redox modifications (D'Autreaux and Toledano, 2007; Heppner et al., 2018; Su et al., 2019). However, the consequence of redox modification on proteins varies depending on both protein itself (e.g., biochemical properties and three-dimensional arrangement) and reactive oxygen species (ROS) (e.g., abundance and species) (Corcoran and Cotter, 2013), which introduces complexes and uncertainties to redox proteome.

Reactive oxygen species are well known as by-products of metabolism and internal source of oxidants in cells (Corcoran and Cotter, 2013) and are harmful to biomolecules if they are overaccumulated, which can eventually be toxic to cells, tissues, and organisms. As another aspect of the two-side sword, ROSs, such as hydrogen peroxide, have been demonstrated as mediators in redox signaling pathways by activating redox modifications to particular target molecules (Bae et al., 2011). Of all the underlying mechanisms that ROS mediates redox signaling, reversible oxidation of proteins with cysteine residues, known as thiol proteins, is well established to trigger and transmit redox signals, which are related to many critical biological activities such as proliferation (Darnell et al., 1994), stress response (Sies, 2014, 2017), and apoptosis (Tobiume et al., 2001; Giorgio et al., 2005; Thangima Zannat et al., 2011).

Redox modifications are widely present in eukaryotes and prokaryotes and involve wildly in multiple biological process and signal pathways. Quantified investigations on the thiol redox proteome have been studied on bacteria (*Escherichia coli*) (Xie et al., 2019), algae (*Chlamydomonas reinhardtii*, *Phaeodactylum tricornutum*, cyanobacterium *Synechocystis*, cyanobacterium *Prochlorococcus*) (McDonagh et al., 2012; Guo et al., 2014; Rosenwasser et al., 2014; McConnell et al., 2018), nematodes (*Caenorhabditis elegans*) (Kumsta et al., 2011), and plants (*Arabidopsis thaliana, Triticum aestivum*) (Bykova et al., 2011; Liu et al., 2014). However, the redox modification of proteome in fungi has not yet been identified so far, so that it is interesting to know about the extent and function of redox in fungi.

Magnaporthe oryzae is a filamentous fungus that causes rice blast, one of the most destructive diseases of cultivated rice and threatens worldwide food production (Wilson and Talbot, 2009; Dean et al., 2012). *M. oryzae* is a model fungus that is used to study the physiological and pathogenic molecular mechanism of the plant pathogenic fungi (Wilson and Talbot, 2009). It has been reported that intricate physiological redox balance is essential for the pathogenicity of *M. oryzae*, such as functional appressorium formation and penetration (Kou et al., 2019). In order to determine roles of the redox modification in fungi, in this study, the thiol-oxidized proteome analysis was performed in *M. oryzae*, through a mass spectrometry (MS)-based strategy. A global landscape of the redox modification sites on protein was shown, providing insights into functions of redox modification in the pathogenic fungi, which also built a groundwork for redox modification studies in other pathogenic or non-pathogenic fungi.

MATERIALS AND METHODS

Strains and Cultural Conditions

The wild-type strain of *M. oryzae* P131 used in this study was maintained on Oatmeal Tomato Agar (OTA) plates at 28°C. Mycelia were incubated in liquid CM cultured on a rotary shaker (180 rpm) for 36 h at 28°C. Conidiation, appressorium formation, and inoculation on barley were performed as described previously (Chen et al., 2014, 2020).

Preparation of Protein Samples

A mixture of mycelia, conidia, appressorium, infection hyphae, and mycelia treated with H_2O_2 was used for protein extraction and liquid chromatography–tandem mass spectrometry (LC-MS/MS) Analysis. Mycelia samples were harvested from liquid CM after shaking culture for 36 h. For H_2O_2-shocked mycelia samples collection, 10 μM H_2O_2 was added to the CM medium after the mycelia was shaking cultured in liquid CM for 36 h, and then, the mycelia were cultured for another 12 h for harvest. Conidia samples were collected after conidiation as described above. Appressorium sample collection was performed as described previously (Chen et al., 2020). For infection hyphae collection, conidia suspension (1×10^5 conidia/mL) was sprayed on the lower leaves of 1-week-old barley and then incubated in a dark, moist chamber at 28°C. The lower barley epidermis together with the infection hyphae on the epidermis were torn down after 48 hpi. All the samples described above were stored at −80°C before processing. Lysis buffer (8 M urea, 1% Triton-100, 10 mM dithiothreitol, and 1% protease inhibitor cocktail) was added to the grinded samples, followed by sonication, refrigerated centrifugation, precipitation, and washing. The protein was redissolved in 8 M urea and quantified with bicinchoninic acid (BCA) method.

High-Performance Liquid Chromatography Tandem Mass Spectrometry Analysis

IodoTMT (Thermo Fisher Scientific, United States) Labeling

Six volumes of cold acetone per 100 μg protein was added to the protein samples followed by precipitation at −20°C for 4 h. Protein was redissolved with 100 μl HES buffer [50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 8.0, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.2% sodium dodecyl sulfate (SDS)] and processed according to the manufacturer's protocol for

iodoTMT kit. Briefly, tris(2-carboxyethyl)phosphine (TCEP) was added to the sample to 10 mM and incubated at 37℃ for 1 h and then mixed with the labeling reagent, which dissolved in methyl alcohol. After incubated in darkness for 1 h at 37℃, six volumes of cold acetone was added to the mixture followed by precipitation at –20℃ for 4 h.

Protein Trypsin Digestion

Trypsin was added at 1:50 trypsin-to-protein mass ratio for the first digestion overnight and at 1:100 trypsin-to-protein mass ratio for a second 4 h digestion at 37℃. The peptides were then desalted and dried by vacuum freeze drying.

HPLC Fractionation

The tryptic peptides were fractionated into fractions by high pH reverse-phase high-performance liquid chromatography (HPLC) using Thermo Betasil C18 column (5 μm particles, 10 mm ID, 250 mm length) (Thermo Fisher Scientific, United States). Briefly, peptides were first separated with a gradient of 8–32% acetonitrile (pH 9.0) over 60 min into 60 fractions. Then, the peptides were combined into four fractions and dried by vacuum freeze drying.

Enrichment

Peptides dissolved in IP buffer (100 mM NaCl, 1 mM EDTA, 50 mM Tris–HCl, 0.5% NP-40, pH 8.0) were incubated with prewashed Anti-TMT antibody beads (Lot number Prod#90076, Thermo Fisher Scientific, United States) at 4°C overnight with gentle shaking. Then, the beads were washed four times with IP buffer and twice with deionized water. The bound peptides were eluted from the beads with 0.1% trifluoroacetic acid, and the eluted fractions were combined and vacuum freeze dried. The resulting peptides were desalted with C18 ZipTips (Millipore, United States).

LC-MS/MS Analysis

The peptides were dissolved in solvent A (0.1% formic acid and 2% acetonitrile) and loaded onto a reversed-phase analytical column. The gradient was comprised of an increase from 8 to 25% solvent B (0.1% formic acid and 90% acetonitrile) over 24 min, 25 to 35% in 8 min and climbing to 80% in 4 min, then holding at 80% for the last 4 min, all at a constant flowrate of 500 nl/min on an EASY-nLC 1000 UPLC system (Thermo Fisher Scientific, United States).

The peptides were subjected to NSI source followed by MS/MS in Q ExactiveTM Plus (Thermo Fisher Scientific, United States) coupled online to the ultraperformance liquid chromatography (UPLC). The electrospray voltage applied was 2.1 kV. The m/z scan range was 350–1800 for full scan, and intact peptides were detected in the Orbitrap at a resolution of 70,000. Peptides were then selected for MS/MS using NCE setting as 28, and the fragments were detected in the Orbitrap at a resolution of 17,500, a data-dependent procedure that alternated between one MS scan followed by 20 MS/MS scans with 15.0 s dynamic exclusion. Automatic gain control (AGC) was set at 5E4. Fixed first mass was set as 100 m/z.

Bioinformatic Analysis

Search Protein Sequence Against Database

We processed all the raw MS data using MaxQuant software (version 1.5.2.8) (Tyanova et al., 2016). The parameters were set as the following. *M. oryzae* strain 70-15 including 12,991 sequences were use as the reference database including two decoy databases (an artifact database to calculate the false discovery rates caused by random matching and a contamination protein database to remove adulterate proteins). Corresponding to the enzyme digestion procedure, Trypsin/P was chosen as cleavage enzyme with up to four missing cleavages. In the two search scenarios, mass tolerance for precursor ions of First search and Main search were adapted to 20 and 5 ppm, respectively, while we set the mass tolerance for fragment ions as 0.02 Da. As to modification, carbamidomethyl on cysteine was considered as fixed modification, while alkylation on cysteine residues, oxidation on methionine residues, and n-terminal acetylation were accounted as variable modifications. We use FDR < 0.01 as adjusted significant threshold, and the minimum score for modification peptides was 40.

Sequence and Motif Analysis

The motif analysis was conducted by MoMo (version 5.0.2) (Cheng et al., 2019) application. A 21-AA window center to the modification sites was use to scan motifs. All the rest parameters were default except that the minimum occurrences was set to 20.

Protein Domain Analysis

Redox-modified proteins (RMPs) were annotated by InterProScan (Quevillon et al., 2005), an online tool based on sequence alignment method, against InterPro domain database.

Function Annotation and Enrichment Analysis

In GO analysis, the identified RMPs were mapped to UniProt-GOA (Huntley et al., 2015) database to get the corresponding UniProt IDs; then, the UniProt IDs were mapped to GO IDs. For RMPs without a UniProt annotation, we used InterProScan to assign GO terms to the proteins. After pooling both set of annotations, the proteins were classified by three GO categories: biological process, cellular component, and molecular function. In Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, we used KASS (KEGG online service tools) to annotate RMPs, and the outcome annotations were mapped to KEGG pathway database by another online tool KEGG mapper. The subcellular localization of RMPs was conducted by the online tool wolfpsort (Horton et al., 2007). The enrichment of assigned functions was estimated using Fisher's exact test with all proteins in *M. oryzae* as background, and the enriched cluster with $p < 0.05$ was considered as significant.

Protein–Protein Interaction Analysis

We selected 425 proteins with functional annotations (i.e., excluding uncharacterized or hypothetical proteins) and stringent quality control thresholds [Maxquant score > 80, $-1.5 \leq$ mass error (ppm) ≤ 1.5] to construct protein–protein interaction (PPI) network for proteins with redox modifications. The PPI network was obtained from STRING (version 11)

(Szklarczyk et al., 2019) database with "confidence score" > 0.7 (high confidence) "active interaction sources" limited to Experiments, Database, Coexpression, and Neighborhood. The figure visualizing the network was constructed by Cytoscape (version 3.8.0) with plugin stringApp (version 1.5.1). To achieve a high confident PPI network, the final confidence score threshold of the PPI network was set to 0.9.

RESULTS

Identification of Thiol Proteome in *M. oryzae*

To identify the redox sites in *M. oryzae*, we performed the proteomic survey of thiol using LC-MS/MS analysis in combination with immunoaffinity enrichment. In order to identify as many redox peptides as possible, we gathered samples from mycelia, conidia, appressorium, infection hyphae of *M. oryzae*, as well as mycelia shocked with reactive oxygen species. The experimental strategy is depicted in **Figure 1A**. In the first repeat, 3273 sites from 1727 proteins were identified while 3299 sites from 1723 proteins in the second repeat. In total, 3713 redox sites belonging to 1899 proteins have been identified, while 2859 redox sites from 1551 proteins were identified by both two repeats, as shown in **Figure 1B** and **Supplementary Table 1**. This result indicated that near 15% of the predicted *M. oryzae* proteins (around 13,000) can be modified by thiol redox, in at least one of the testing samples. The identified redox sites and the corresponding proteins were analyzed from different aspects to shed a light to the redox proteome in *M. oryzae*.

Patterns of the Redox Modified Sites and Proteins

Among the identified thiol redox-modified proteins, around 50% had only one thiol-modified site, and 22, 10, or 3.2% had two, three, or four thiol modified sites, respectively. Other proteins had more than four thiol-modified sites, even up to 22 sites (**Figure 1C**). To discover the common features shared by proteins containing redox modification sites, amino acids motifs besides the modification sites were identified. The normal pattern of the motifs could indicate important function hub in respect to evolutionary conservation; therefore, we calculated the abundance of amino acids around the redox modification sites. The redox modification-specific motifs were discovered by MoMo Tools, which is a particular effort to characterize post-translational modifications in proteins. The detected motifs show the amino acids abundance around the redox modification sites (**Figure 1D**). By the abundance analysis, we found both positively

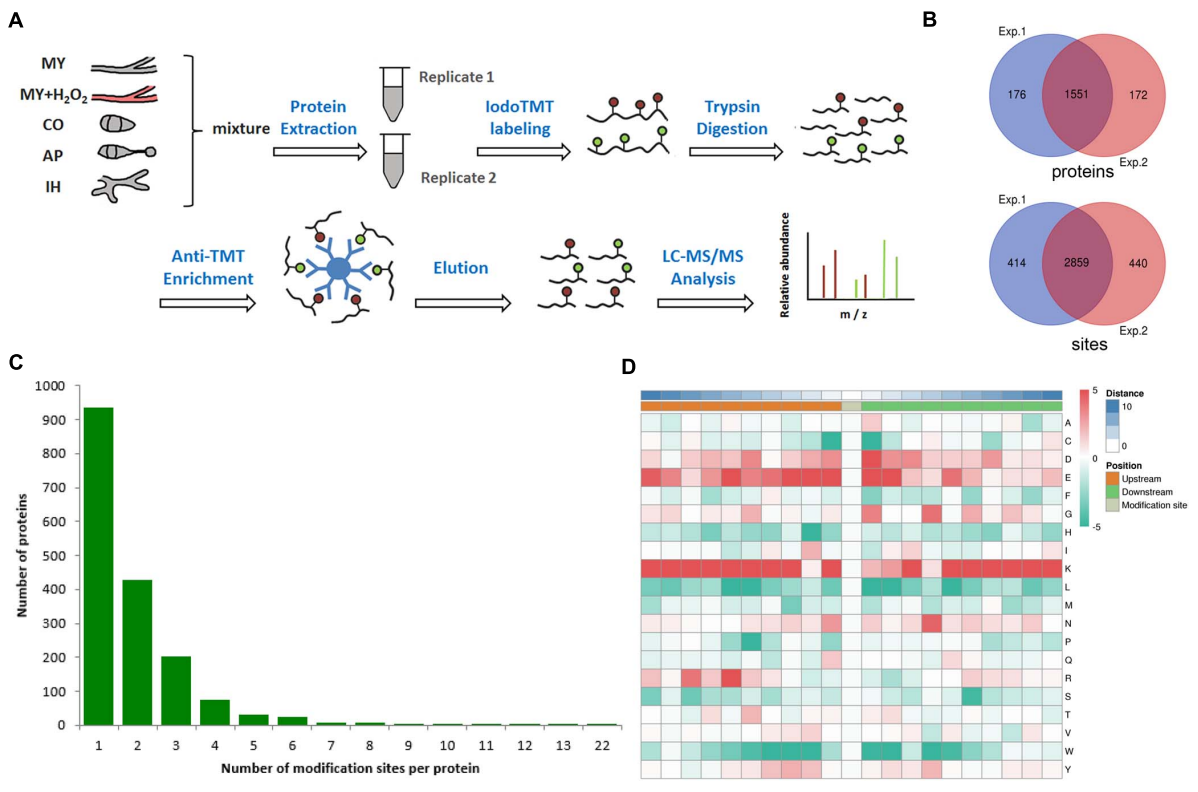


FIGURE 1 | Redox proteomics identification of thiol-modified proteins in *M. oryzae*. **(A)** Flow chart of the integrated strategy for quantitative analysis of the redox proteome. **(B)** Venn diagram of modified proteins identified in replicates. **(C)** Venn diagram of modified sites identified in replicates. **(D)** The motif enrichment heatmap of upstream and downstream amino acids of all identified modification sites. Red and green indicate that this amino acid is more and less abundant close to the modification site.

and negatively charged amino acids (i.e., lysine and arginine; glutamic acid and aspartic acid) overrepresented in a 21-residues wide window centered to the redox modification site. However, when it comes to motifs, only positively charged residues are examined as principle conserved elements. In addition, the residues are mainly located in the upstream, which is at least three residues away from the modification sites. These findings suggest that redox reactive modification actively happens to lysine-rich segments and basic polar amino acids regions (**Figure 1D**). The corresponding three-dimensional protein structures and biochemical properties derived from the motifs could manipulate the sensitivity of thiols/disulfides in *M. oryzae*.

Protein domains, by definition, are conservative amino acids sequences, the length of which vary between 25 and 500 amino acids, shared by various protein molecules repeatedly. The shared sequences are basic evolutionary components and usually function and structure similarly. To explore the evolutionary conservation of all identified RMPs, we estimated the enrichment of protein domains. The top 5 enriched domains are "Tim10/DDP family zinc finger" (Jarosch et al., 1996; Paschen et al., 2000), "Gelsolin repeat" (Silacci et al., 2004), "Dnaj central domain" (Qiu et al., 2006), "CRAL/TRIO, N-terminal domain" (Panagabko et al., 2003), and "CRAL/TRIO domain" (Panagabko et al., 2003; **Figure 2A**). In brief, the leading domains function as switcher that control mitochondrial protein import (Jarosch et al., 1996; Paschen et al., 2000), as organizer that control actin arrangement (Silacci et al., 2004), as chaperones that protect proteins (Qiu et al., 2006), and as modulators that mediate interaction between retinoids and visual cycle enzyme (Panagabko et al., 2003).

The translocation of a protein determines the exact locus in which the protein fulfills its mission consequently; therefore, localization of proteins at the subcellular level could benefit, for example, the identification of drug targets. However, the aberration of protein translocation endangers the homeostasis of organisms. Moreover, the controlled distribution of redox proteome is pivotal to intracellular physiological balance; for instance, redox processes are well regulated in mitochondria to protect cellular components (Rigoulet et al., 2011). Therefore, the generalized subcellular localization of RMPs contributes to our understanding on redox proteome in *M. oryzae*. To give an insight into the distribution of RMPs at the subcellular level, we predicted the subcellular localization of identified redox proteins and further classified them based on the assigned loci. In **Figure 2B**, the identified proteins distributed in different locations and the major four predictions are cytoplasm (27.38%), mitochondria (20.48%), nucleus (19.8%), and extracellular (17.64%), suggesting that the redox-modified proteins are widely distributed and play distinct roles in different subcellular regions of the cell.

Thiol Redox Modification Is Involved in Many Cellular Processes and Metabolisms

For the goal to disclose how proteins undergoing redox reaction contribute to the various biological processes in *M. oryzae*, proteins identified with redox modification sites in this study were annotated by Gene Ontology (GO). Considering the informative difference across GO categories, level 2 GO annotations were chosen to represent features and characteristics of proteins that potentially perform important roles in different biological processes (BP), cellular components (CC), and molecular functions (MF) (**Supplementary Table 2** and **Figures 3A–C**).

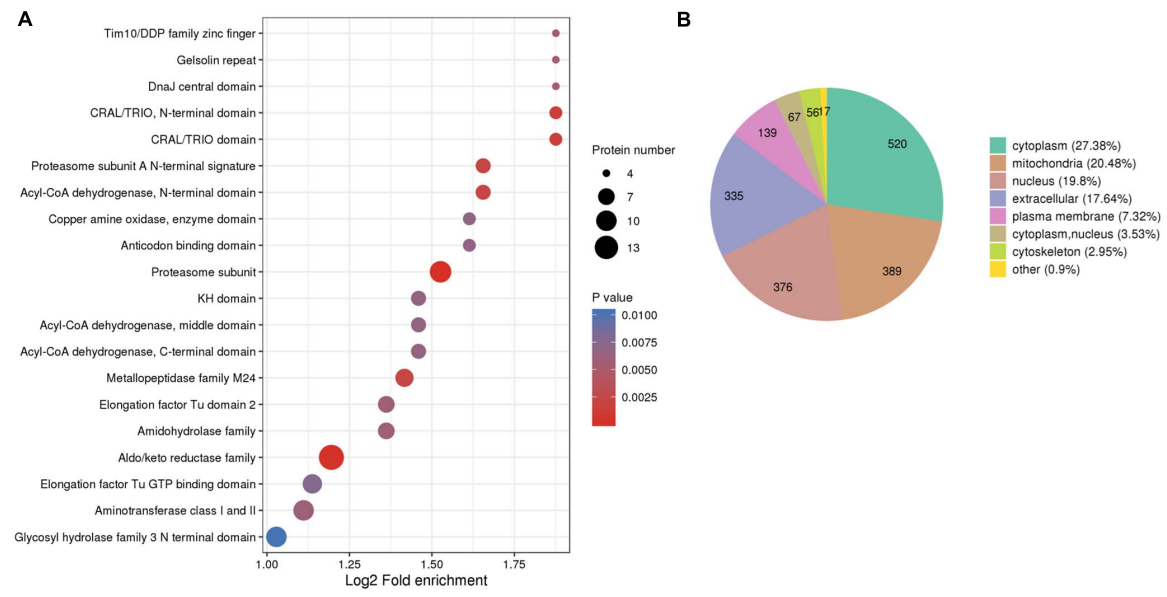


FIGURE 2 | Protein domain enrichment and subcellular localization of the identified thiol proteins. **(A)** Protein domain enrichment bubble plot of proteins corresponding to modification sites. **(B)** Subcellular localization chart of proteins corresponding to modification sites.

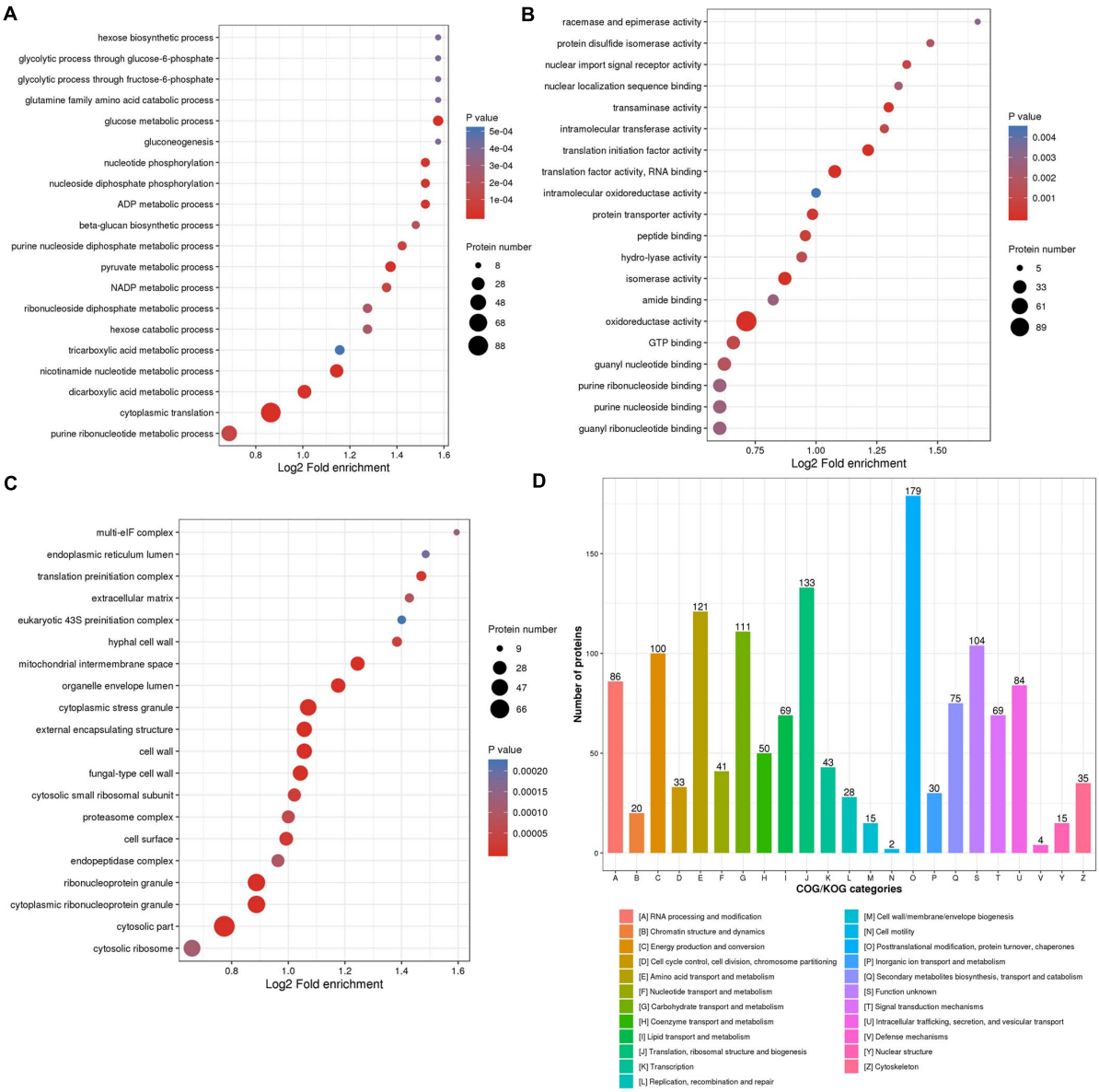


FIGURE 3 | Enrichment categories of the identified thiol proteins. Gene Ontology (GO) classification enrichment bubble plots of the thiol proteins were shown based on **(A)** biological processes, **(B)** cellular component, and **(C)** molecular functions. **(D)** Clusters of orthologous groups functional classification chart of the thiol proteins.

All of the identified RMPs were assigned level 2 GO terms using the KEGG mapper (Kanehisa and Sato, 2020), and an enrichment analysis was performed to exhibit the overrepresentation of terms (including but not limited to the level 2) related to redox proteome in *M. oryzae*. As shown in **Figure 3A**, regarding biological processes, RMPs are mostly enriched in nutrient utilization processes, including hexose biosynthetic process, glycolytic processes through glucose-6-phosphate and fructose-6-phosphate, glutamine family amino acid metabolic process, glucose metabolic process, hexose metabolic process, gluconeogenesis, as well as tricarboxylic acid metabolic process. This result suggested that redox modification plays key roles in maintaining cellular processes for fungal survival and development. Many other cellular processes such as nucleotide phosphorylation, ADP and nicotinamide adenine dinucleotide phosphate (NADP) metabolic processes, beta-glucan biosynthetic process, purine and pyruvate metabolic processes, cytoplasmic translation, etc. were also significantly enriched (**Figure 3A**), indicating widespread functions of redox modification in cellular processes.

The mostly enriched categories in molecular functions included racemase and epimerase activity, protein disulfide isomerase activity, intramolecular oxidoreductase activity, amide binding, oxidoreductase activity, and GTP binding activity, which were well consistent with the original function of redox modification (**Figure 3B**). Besides, the redox modification

may also affect activities of nuclear import signal receptor, nuclear localization sequence binding, intracellular transferase, translational initiation factor, RNA binding, protein transporter, peptide binding, hydrolyase, isomerase, guanyl nucleotide and ribonucleotide binding, and purine nucleoside and ribonucleoside binding (**Figure 3B**), also suggesting a widespread role of the redox modification in regulating protein functions.

Categories belonging to cellular components were mostly enriched in multi-eIF complex, which reflects its role of oxidoreduction (**Figure 3C**). The translation-related complexes, such as endoplasmic reticulum lumen, translation preinitiation complex, eukaryotic 43S preinitiation complex, cytosolic small ribosomal subunit, and cytosolic ribosome were also significantly enriched, suggesting a regulatory role of redox modification in translation (**Figure 3C**). Other enriched cellular components included cell wall and extracellular component (extracellular matrix, hyphal cell wall, external encapsulating structure, cell surface, etc.), cytoplasmic stress granule, organelle envelope lumen, proteasome complex, endopeptidase complex, and ribonucleoprotein granule (**Figure 3C**).

In sum, GO annotations could demonstrate the common features across the whole redox proteome and shed a light on the underlying mechanisms of redox-related biological processes and molecular mechanisms in *M. oryzae*.

To understand the function of identified RMPs, we also grouped them into Clusters of Orthologous Groups of Proteins (COG) (**Supplementary Table 2**; Tatusov et al., 2000). The orthologous groups are known as conserved clusters sharing identical or analogous biological functions and biochemical properties. As shown in **Figure 3D**, the function clusters are mainly associated with the following: (i) (post-)translation and (post-)transcription including "post-translational modification, protein turnover, chaperones," "translation, ribosomal structure and biogenesis," "RNA processing and modification;" (ii) metabolism such as "amino acid transport and metabolism," "carbohydrate transport and metabolism," and "energy production and conversion." These results were well consistent with GO annotation of biological processes, cellular components, and molecular functions.

Metabolic Pathway Analysis of RMPs

Biological pathways are functionally united sets of genes that are proven to contribute in specific biological processes, such as energy metabolism. Hence, mapping RMP genes to pathways is an ideal approach to illustrate the roles of RMPs in redox proteome. We mapped RMPs to pathways in KEGG and enriched mapped pathways to explore how RMPs contribute to specific biological processes (**Supplementary Table 2**). We found that the top 5 significant pathways that RMPs are gathered in are "carbon metabolism" (mgr01200; FDR, 1.17E-6; fold enrichment, 1.74), "biosynthesis of secondary metabolites" (mgr01110; FDR, 1.09E-7; fold enrichment, 1.31), "alanine, aspartate and glutamate metabolism" (mgr00250; FDR, 2.37E-7; fold enrichment, 2.31), "biosynthesis of amino acids" (mgr01230; FDR, 1.92E-5; fold enrichment, 1.50), and "valine, leucine and isoleucine degradation" (mgr00280; FDR, 1.41E-3; fold enrichment, 1.89) (**Figure 4A**). The top 20 significant pathways included primary metabolisms of carbon metabolism, starch and sucrose metabolism, glycolysis/gluconeogenesis, citrate cycle, galactose metabolism, and pentose phosphate pathway; biosynthesis of secondary metabolites metabolism; purine, pyruvate, and pyrimidine metabolisms; and amino acids metabolism (**Figure 4A**). These results suggest the wide range functions of redox modifications in primary metabolism, secondary metabolism, as well as nucleotide and amino acid metabolisms of *M. oryzae*.

Protein Interaction Networks of RMPs

Protein–protein interactions contribute to understanding on cell physiology states, as they play essential roles in almost every process, for instance, signaling transduction. To find interactions among RMPs in *M. oryzae*, a PPI network was built by screening 425 identified RMPs with stringent parameters (**Figure 4B**). Interestingly, by coloring proteins based on function annotations, we found intricate linkages among RMPs. The important network notes include ribosome, translation initiation, aminoacyl-tRNA biosynthesis, protein processing in ER, and proteasome, indicating that the proteins involved in translation and degradation were enriched. Network notes also include heat shock protein, chaperone, and mitogen-activated protein kinase (MAPK) signaling pathway, suggesting important roles of RMPs in stress and signaling responses. The T-complex protein note was also enriched (**Figure 4B**). Considering that redox modifications of proteins are important PTMs that affect protein functions, the complicated interactions revealed by the PPI network boost the variety and complexity of biological activities of *M. oryzae*.

Redox Modified Proteins Involved in Development, Stress Response, and Pathogenicity

In order to determine functions of redox modification in fungal pathogenesis, we searched the reported *M. oryzae* development, stress response, and pathogenesis-related proteins using all identified RMPs. Notably, as many as 174 proteins that were previously reported as pathogenesis-related proteins were identified to be thiol oxidized in our study. These proteins are associated with a wide range of pathogenesis-related cellular functions, involved in crucial biological process, such as autophagy, ubiquitination, and glycogen metabolism; crucial infection-related pathways, such as MPG1, MagB, GTPases, SLN1, PMK1-MAPK, and MPS1; important subcellular structures, such as septin ring, mitochondria, cytoskeleton, cell wall, and proteasome components; and vital proteins, such as protein kinases and phosphatases. Most of the proteins contained less than three redox sites, while a few also contained more than seven sites, implying a deep involvement of redox modification in pathogenic processes. The complete list of identified redox-modified development or pathogenicity-related proteins is provided in **Supplementary Table 3**. Using information of these reported thiol redox proteins, we are able to build a model that shows predicted roles of redox modification in the development and infection process of *M. oryzae* (**Figure 5**). Functions of these

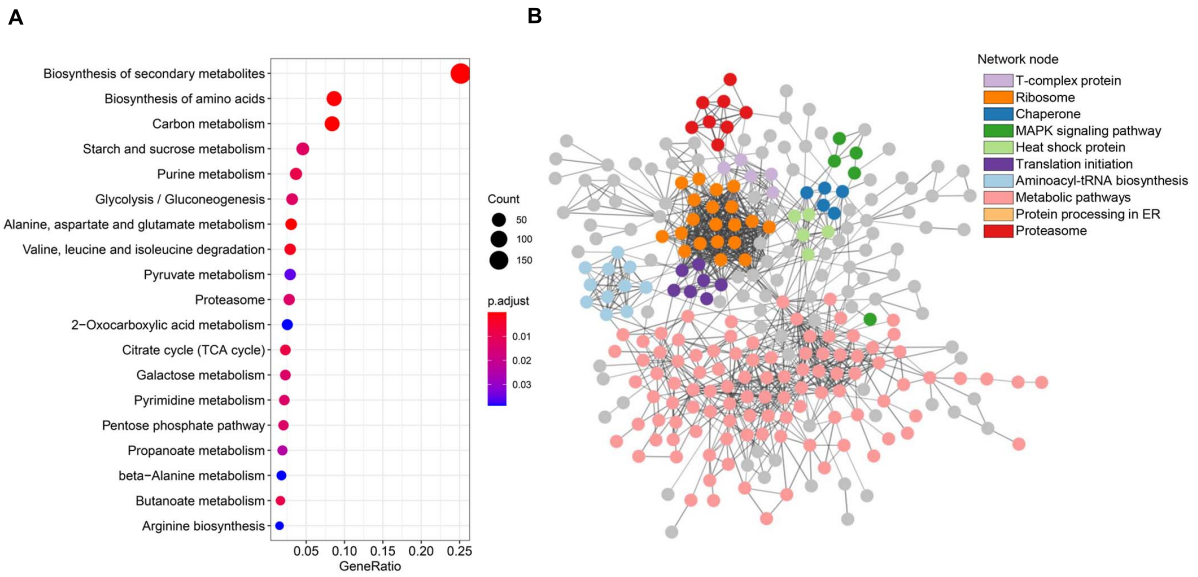


FIGURE 4 | Metabolic pathway and protein–protein interaction analyses of the thiol proteins. **(A)** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment bubble plot of the identified thiol proteins. **(B)** STRING network analysis of identified thiol proteins.

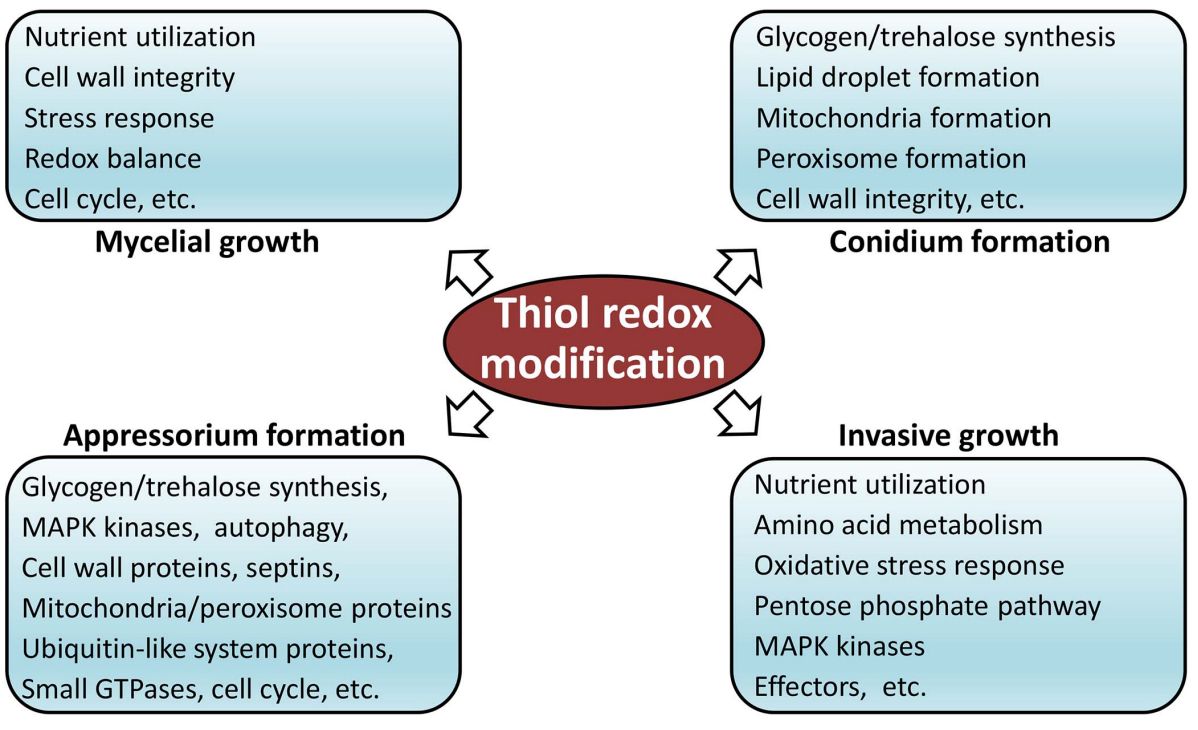


FIGURE 5 | Model of thiol redox-modification-regulated development and infection of *M. oryzae.*

reported redox-modified proteins will be discussed in more detail in the following sections.

Autophagy

There are eight identified proteins, including cysteine protease ATG4, beclin-1 ATG6, ubiquitin-like protein ATG8, lipase ATG15, sorting nexin-4 ATG24, transmembrane protein ATG27, and DNA binding SaNT domain protein MoSnt2, that were previously demonstrated to be related to autophagy. ATG6, a part of a complex with class III phosphatidylinositol T-kinase (PI3K)/Vps34, has been shown to function in vacuolar protein sorting (VPS) and autophagy in yeast, and the conserved domain

also has been identified (Furuya et al., 2005). ATG8 is a member of an evolutionarily conserved ubiquitin-like protein family and can be conjugated to the lipid phosphatidylethanolamine. It is involved in multiple membrane trafficking pathways including autophagy and is required for double membrane-bound structures formation of autophagosomes in *Saccharomyces cerevisiae*. ATG4 is a deconjugation enzyme that can recognize the residues in ATG8 and modulated ATG8 reversibly (Fass et al., 2007; Nakatogawa et al., 2007). ATG15 is required for degradation of autophagic bodies and is essential for the maintenance of lipid droplets amount in the stationary phase in *S. cerevisiae*. The loss of ATG15 leads to enhanced lipolytic activity (Maeda et al., 2015). It has been demonstrated that loss of MoATG24 disrupts mitophagy and consequently results in significantly reduced conidiation. During mitophagy, MoATG24 directly associated with and recruited mitochondria to the autophagic structures (He et al., 2013). ATG27 is required for ATG9 cycling and for specific autophagy in yeast (Yen et al., 2007). MoSnt2 recognizes histone H3 acetylation via its PHD1 domain to recruit the histone deacetylase complex, giving rise to deacetylation of H3 and direct links to MoTor signaling, thus regulating infection-associated autophagy and infection in an epigenetic mechanism. The null mutants of MoSNT2 are impaired in various aspects including infection structure development, conidiation, oxidative stress tolerance, cell wall integrity, autophagy-dependent fungal cell death, and pathogenicity and are compromised in autophagy homeostasis (He et al., 2018).

Glycogen Metabolism

Four oxidized proteins, including glycogen debranching enzyme AGL1, glycogen phosphorylase GPH1, glycogen synthase GSN1, alpha, alpha-trehalose-phosphate synthase 1 TPS1, and trehalase TRE1, were reportedly related to glycogen metabolism. AGL1 and GPH1 are required for mobilization of glycogen stores during appressorium development in *M. oryzae*. It has been demonstrated that the deletion mutants of AGL1 and GPH1 showed significant reduction in pathogenicity and in expression of TPS1 and TPS3. TPS1 regulates virulence-associated gene expression in association with Nmr transcriptional inhibitors, by responding to glucose-6-phosphate levels and the balance of NADP/NADPH. GPH1 plays roles in the breakdown of glycogen and is essential in mobilizing glycogen and full virulence. The null mutants of GSN1 significantly reduced the synthesis of intracellular glycogen (Badaruddin et al., 2013). TRE1 locates in the cell wall with characteristics of both neutral and acidic trehalases and is dispensable for pathogenicity, but trehalose degradation is important for the efficient development in plant tissue following initial infection (Foster et al., 2003).

Ubiquitin-Like Modification

Seven well-reported ubiquitination-related proteins, including E3 ubiquitin-protein ligase MoBRE1 and MoUBR1, SCF E3 ubiquitin ligase complex F-box protein MoGrr1, E3 ubiquitin ligase complex SCF subunit scon-3 MoSkp1, SUMO protein MoSmt3, SUMO-activating enzyme subunit MoUba2, and ubiquitin carboxyl-terminal hydrolase UBP14, were observed to be redox modified. MoBre1 was shown to be required for growth, conidiation, and pathogenicity in *M. oryzae*, in which its loss resulted in the reduction in di- and tri-methylation level of histone 3 lysine 4 (H3K4). MoUbr1 is essential for conidial adhesion and germination, as it degrades components of cAMP/PKA and MAPK Pmk1 signaling pathways by the N-end rule pathway (Shi et al., 2016). The *Mogrr1* null mutants displayed defects in vegetative growth, melanin pigmentation, conidial production, and resistance to oxidative stress and full virulence and could not generate sufficient turgor pressure in the appressorium for the penetration into plant tissues. The expression of central components of the MAP kinase and cAMP signaling pathways, which are required for appressorium differentiation, was reduced in Mogrr1 mutants (Guo et al., 2015). Skp1 is part of the Skp1-Cullin 1-F-box (SCF) E3 ubiquitin ligase complex, which is necessary for protein degradation, and MoSkp1 is located in spores and germ tubes, with abundant expression in appressoria. The reduction in MoSkp1 results in the decrease in total protein ubiquitination and, thus, defective cell cycle and appressorial development (Prakash et al., 2016). MoSmt3 is a small ubiquitin-related modifier (SUMO) protein present in both the nucleus and cytoplasm, while MoUBA2 is an E1 enzyme present in the nucleus. Nevertheless, both MoSmt3 and MoUba2 are in the nucleus under oxidative stress conditions. Deletion mutants of *MoSMT3* and *MoUBA2* showed significant defects in mycelial growth, conidiation, septum formation, conidial germination, appressorium formation, and pathogenicity, suggesting that they are crucial for infection-related development, stress responses, and pathogenicity in *M. oryzae* (Lim et al., 2018). It has been demonstrated that the ortholog of MoUbp14 plays general roles in ubiquitin-mediated protein degradation in *S. cerevisiae*. Destruction of *MoUBP14* in *M. oryzae* led to loss of pathogenicity and severe reduction in mycelial growth, sporulation, carbon source utilization, and increasement in sensitivity to distinct stresses (Wang et al., 2018).

Crucial Infection-Related Pathways

Several well-known proteins in infection-related pathways, such as hydrophobin-like protein MPG1, guanine nucleotide-binding protein subunit alpha MagB, sensor protein SLN1, CMGC/MAPK/ERK protein kinase PMK1, and CMGC/MAPK protein kinase MPS1, were identified as redox-regulated proteins. MPG1 is a small, secreted, cysteine-rich, moderately hydrophobic protein that directs the formation of a rodlet layer on conidia that are interwoven similarly to 5-nm rodlets, thus interacting with hydrophobic surfaces and acting as a developmental sensor for appressorium formation. Deletion mutants of Mpg1 showed reduced efficiency of appressorium formation resulting in impaired pathogenicity (Talbot et al., 1993; Beckerman and Ebbole, 1996; Talbot et al., 1996). MAGB is involved in multiple signal transduction pathways controlling vegetative growth, conidiation, conidium attachment, appressorium formation, mating, and pathogenicity (Liu and Dean, 1997). MoSLN1 acts as a pathogenicity factor that is involved in responses to osmotic stress, the cell wall integrity, and the activity of peroxidases via modulation of intra- and extracellular peroxidase activities. The MoSLN1 mutants displayed hypersensitivity to various stresses,

reduced sensitivity to cell wall perturbing agent calcofluor white, and loss of pathogenicity (Zhang et al., 2010). PMK1 is a member of a highly conserved MAP kinase signal transduction pathway, which acts downstream and collaborate with a cAMP signaling pathway for infection structure formation. Disruption of PMK1 leads to loss of the ability to form appressoria invasively in rice plants (Xu and Hamer, 1996). MPS1 is a mitogen-activated protein kinase that is crucial for pathogen penetration, and the mutant of Mps1 showed sensitivity to cell-wall-digesting enzymes, reduced sporulation, and complete loss of pathogenicity resulting from the inability of appressoria to penetrate plant cell surfaces (Xu et al., 1998).

GTPases

Redox modification was observed in 11 GTPases, including rho-type GTPase-activating protein LRG1, cell division control protein MgCdc42, rab GDP-dissociation inhibitor MoGdi1, arf GTPase-activating protein MoGlo3, hydrophobin-like protein MPG1, rho-GTPase-activating protein MoRga4 and MoRga5, GTP-binding protein MoRab5B and Rho3, and Ras-like protein Rac1 and MoYpt7. LRG1 was reported as key regulators of infection-associated morphogenesis of *M. oryzae*. Deletion mutants of LRG1 presented significant reduction in hyphal growth, loss of pathogenicity, and decreased sensitivity to cell-wall-perturbing agents (Li et al., 2014). Cdc42 is a member of the Rho-family small GTP-binding proteins and function as a pivotal signaling switch that controls actin cytoskeleton organization and cell polarity, cycling between active GTP-bound and inactive GDP-bound forms. MgCdc42 has been shown to be essential for plant penetration, and disruption of MgCdc42 leads to dramatically reduced pathogenicity resulting from the arrest of penetration and infectious growth due to the defect of turgor and superoxide generation during the appressoria development. Furthermore, the MgCdc42 mutants showed pleotropic defects including gherkin-shaped conidia, delayed germination, and decreased sporulation (Zheng et al., 2009). MoGlo3 is highly expressed during conidiation and early infection stages, which is required for endocytosis, reactive oxygen species scavenging, endoplasmic reticulum (ER) stress response, vegetative growth, conidial production, sexual development, appressorium function, and pathogenicity (Zhang et al., 2017). MoRab5B has been demonstrated to be crucial for vegetative growth and development, conidiogenesis, melanin synthesis, vacuole fusion, endocytosis, sexual reproduction, and plant pathogenesis in *M. oryzae* (Yang et al., 2017). Rac1 takes part in actin cytoskeleton organization and polarized cell growth in many organisms. MgRac1 is required for normal conidial production in amount and morphology, as well as crucial for appressorial formation and pathogenicity (Chen et al., 2008). Rho3 is essential for plant infection and acts as a critical regulator of developmental processes and pathogenicity. Appressoria formed in the absence of Rho3 showed morphological abnormality and defect in plant penetration. Overexpression of MgRho3 contributes to the infectivity of *M. oryzae* (Zheng et al., 2007). Previous study has showed that MoYpt7 is involved in fungal morphogenesis, vacuole fusion, autophagy, stress resistance, and pathogenicity in *M. oryzae*, and the disruption mutants of MoYpt7 exhibited impairment in autophagy, breached cell wall integrity, higher sensitivity to both calcium and heavy metal stress, malformed conidia, defection in appressoria formation, and the inability to cause disease (Liu et al., 2015).

Protein Kinases

Ten pathogenesis-related protein kinases were identified as thiol oxidized, including CAMK/CAMK1 protein kinase MoCMK1, CMGC/GSK protein kinase MoGsk1, CMGC/MAPK/P38 protein kinase MoHog1, CMGC/MAPK protein kinase MPS1, CMGC/MAPK/ERK protein kinase PMK1, CAMK/CAMKL/AMPK protein kinase MoSNF1, guanylate kinase MoGuk2, STE/STE7 protein kinase MoMkk1, STE/STE11 protein kinase MST11, and serine/threonine-protein phosphatase MoPpe1. MoCMK1 was shown playing key roles in the pathogenicity of *M. oryzae*, and the MoCMK1 mutants had sparse aerial hyphae, fewer conidia, and a delay in conidial germination and appressorial formation (Liu et al., 2010). The expression of MoGSK1 is regulated by MPS1, another redox-modified MAP kinase, and the deletion of MoGSK1 leads to severe delay in mycelial growth, complete loss of conidiation, and inability to penetrate the host surface by mycelia-formed appressorium-like structures, thus losing the pathogenicity (Zhou et al., 2017). MoSNF1 is critical for sporulation, vegetative growth, and pathogenicity of *M. oryzae*. The MoSNF1 mutant produces less and abnormal conidia and was largely impaired in conidial germination and appressorium formation (Yi et al., 2008). MoGuk2 is involved in the *de novo* GTP biosynthesis pathway and is important for infection-related morphogenesis in *M. oryzae* (Cai et al., 2017). MoMkk1 plays a role in modulating intracellular cAMP levels and response to osmotic stress, and disruption of MoMKK1 resulted in less aerial hypha, defective asexual development, and reduced pathogenicity (Yin et al., 2016). MST11 collaborates with MST7 and PMK1 as a MAP kinase cascade to regulate infection-related morphogenesis (Zhao et al., 2005). MoPpe1 takes part in linking CWI and TOR signaling and is essential for vegetative growth, conidiation, and full virulence (Qian et al., 2018).

Subcellular Structures

Proteins located on subcellular structures play various and crucial roles in hyphal growth, conidiogenesis, appressorium development, and pathogenesis of *M. oryzae*, such as cell-wall-related proteins, mitochondrial proteins, and septin proteins. Our data showed that 16 cell-wall-related proteins were redox modified, including chitin deacetylases CDA1, glycosyl hydrolase CDA3, 1,3-beta-glucan synthase component FKS1, alpha-1,3-glucan synthase MoAGS1, tetrahydroxynaphthalene reductase MoBUF1, cell wall biogenesis protein phosphatase SSD1, beta-1,6-galactanase MC63, neutral alpha-glucosidase MoGls2, chitin synthase MgCHS1, MgCHS4, MgCHS5, MgCHS6, and 1,3-beta-glucanosyltransferase Gel1, Gel2, Gel3, and Gel4. Nine mitochondrial proteins were thiol oxidized, including enoyl-CoA hydratase Ech1, peroxisomal hydratase-dehydrogenase-epimerase MFP1, alanine-glyoxylate aminotransferase AGT1, dynamin-A MoDnm1, peroxisome

biosynthesis protein MoPEX1, C-8 sterol isomerase ERG2, and D-lactate dehydrogenase MoDLD1, MoDLD2, and MoDLD3. Three septin proteins, eptin-like spn2 MoSep4 and cell division control proteins MoSep5 and MoSep6, were redox regulated. These results revealed the broad roles of redox modification in pathogenicity regulation of *M. oryzae.*

Effectors

Effectors play key roles in fungal–plant interactions during fungal invasive growth. Therefore, we also screened the predicted effector proteins according to the previous reports (Ref). In total, 72 thiol-modified sites in 39 putative effectors were identified (**Supplementary Table 4**). These include several reported effector/elicitor proteins, MoHEG16, MoHrip2, and MoMSP1 (**Supplementary Table 3**). MoHEG16 is highly expressed in early stage of the invasive hyphae and required full virulence (Mogga et al., 2016). MoHrip2 is an apoplastic effector, which suppresses host immunity and is also required for the full virulence of *M. oryzae* (Nie et al., 2019). The snodprot1 protein MSP1 is another virulence factor in *M. oryzae* (Jeong et al., 2007). These data showed that the effector proteins were also commonly modified by redox modification, suggesting that redox modification could regulate functions of effector proteins for invasion in host cells.

DISCUSSION

In this study, we systematically explored the landscape of RMPs in *M. oryzae*, which is the first whole redox proteome analysis for rice blast fungus and for fungal pathogen. In total, 1899 proteins with 3713 redox modification sites were identified and studied with different bioinformatic approaches, including motif analysis, domain identification, subcellular localization, and GO and KEGG pathway annotation with corresponding enrichment. An enrichment of lysine residues was found around the cysteine redox modification sites. Regarding localization, the majority (over two-thirds) of redox modifications on proteins occurs in the cytoplasm, mitochondrion, and nucleus. With respect to function annotation and enrichment, RMPs are highly associated with metabolism (e.g., carbon metabolism) and biosynthesis (e.g., secondary metabolites), which is expected. However, the results also show that protein complex (e.g., proteasome and spliceosome, respectively) that evolved in fundamental processes (translation and transcription) can be affected by redox modification.

Redox modification widely exists in both eukaryotes and prokaryotes. The development of high-specificity labeling and enrichment, as well as high-resolution MS techniques, promotes the global thiol redox proteome studies. As early as 2011, redox-sensitive proteome has been investigated, comparing between dormant and non-dormant seeds of wheat (*T. aestivum*). Altogether, 193 reactive Cys were found in 79 unique proteins. The identified proteins are involved in protein synthesis and storage, carbohydrate metabolism, proteases, transport, and signal transduction, which are associated with seed dormancy and protection (Bykova et al., 2011). The redox thiol proteome of *Caenorhabditis elegans* treated with short-term H_2O_2 stress was also quantified, and 40 different proteins containing oxidation-sensitive cysteines were identified. These proteins play roles in mobility and feeding protein translation and homeostasis and adenosine triphosphate regeneration (Kumsta et al., 2011). Quantifications of redox thiol proteome were performed on algae, and 80 proteins were identified, which revealed the response to lack of nitrogen by cyanobacterium via post-translational redox changes (McDonagh et al., 2012). In the cyanobacterium *Synechocystis*, approximately 2100 Cys sites from 1060 proteins under light, dark, and photosystem II inhibitor-treated conditions were quantified, implying the broad redox regulation of photosynthetic organisms (Guo et al., 2014). In *P. tricornutum*, the degrees of oxidation of 3845 cysteines were quantified, and 278 redox-sensitive proteins were identified to elucidate the redox-sensitive signaling network (Rosenwasser et al., 2014). McConnell et al. combined the reversibly thiol oxidation and the phosphorylation in proteome quantitative analysis to determine the extent of phosphorylation in the redox thiol proteome in *Chlamydomonas reinhardtii*. A total of 3353 oxidized Cys sites on 1457 enriched proteins were identified, which found the possibility of crosstalk between redox modification and other PTMs (McConnell et al., 2018). Regarding plants, the thiol redox proteome was quantified in *A. thaliana* in 2014 to investigate the oxidative stress after H_2O_2 treatment and in 2017 to study the CO_2 response via bicarbonate treatment. One hundred ninety-five cysteine-containing peptides from 179 proteins and 903 cysteine-containing peptides with 47 significantly changed proteins were identified, respectively. These studies revealed the roles of redox modification in oxidative stress response and CO_2 response (Liu et al., 2014; Yin et al., 2017). Recently, the thiol redox proteome was quantified in bacteria, *Escherichia coli*, to assess the thiol oxidation status of *E. coli* after being phagocytized by neutrophils. A total of 173 cysteine containing peptides representing 117 proteins were identified, indicating that neutrophil phagocytosis leads to an overall breakdown of the *E. coli* protein thiol homeostasis, which might be a general antimicrobial mechanism for neutrophils to counteract invading bacteria. However, there are no reports on thiol redox proteome of pathogenic fungus, and our knowledge on the global redox modification in pathogenic fungus is still limited.

In order to identify as many redox proteins as possible, acquiring a comprehensive view of the global thiol redox proteome in *M. oryzae*, we employed the samples from all the development stages including mycelia, conidia, appressoria, and infection hyphae. Furthermore, we generated an oxidative stress condition through H_2O_2 treatment on mycelia. We identified a large number of redox proteins (1899 proteins) from the mixture described above, contrasted with the number of identified proteins in studies on other species, such as *T. aestivum* (79 proteins), *C. elegans* (40 proteins), cyanobacterium *Prochlorococcus* (80 proteins), cyanobacterium *Synechocystis* (1060 proteins), *P. tricornutum* (278 proteins), *C. reinhardtii* (1457 proteins), *A. thaliana* (179 proteins), and *E. coli* (117 proteins).

Notably, among the redox-regulated proteins identified in the present study, up to approximately 10% (174) proteins are previously reported to be pathogenesis related. These proteins are involved in autophagy, ubiquitination, glycogen metabolism, GTPases, MPS1 pathway, PMK1-MAPK pathway, septin ring, cytoskeleton, cell wall integrity, etc. Autophagy is essential for *M. oryzae* infection, by regulating programmed cell death during appressorium maturation (Veneault-Fourrey et al., 2006) or maintaining lipid body integrity (Liu et al., 2007). Autophagy also regulated carbohydrate catabolism and homeostasis spatially and temporally to ensure successful conidiation in *M. oryzae* (Deng et al., 2009; Deng and Naqvi, 2010). The ubiquitin system modulates protein functions through targeting substrates for ubiquitination and coordinates with cell cycle proteins and controls various cell functions (Prakash et al., 2016; Shi et al., 2016; Lim et al., 2018). The mobilization of glycogen and trehalose is required for appressorium development to fuel biosynthetic processes and turgor generation in the developing appressorium (Thines et al., 2000). Some Rho GTPases interact with each other and regulate signal transduction pathways in eukaryotes. Interestingly, the expression of Rho GTPases MgRac1 was negatively regulated by MgCdc42 and MgRho3, respectively, while all of the three proteins are thiol oxidized, implying a deeply interaction in redox regulation (Zheng et al., 2006). The Pkc1-Mps1 MAP kinase cell wall integrity (CWI) signaling pathway is essential for the production and function of *M. oryzae* appressoria. Disruption of MoMPS1, and terminal MAP kinase of cell wall integrity pathway, leads to appressoria that are unable to penetrate the host plant, thus loss of pathogenicity (Xu et al., 1998; Jeon et al., 2008). The cytoskeleton plays roles in the regulation of multiple cellular processes including but not limited to morphogenesis, cytokinesis, establishment of cell polarity, endocytosis, and exocytosis (Moseley and Goode, 2006; Berepiki et al., 2011). Pore formation at the base of the appressorium is required for the reorganization of the F-actin meditated by a septin ring (Dagdas et al., 2012). In conclusion, these results indicate that redox modification plays an important role in the regulation of pathogenicity in *M. oryzae*.

Interestingly, we found that several pathogenesis-related proteins are able to be regulated by more than a single post-translational modification. It has been reported in *C. reinhardtii* that some of the redox-regulated proteins were found to be additionally phosphorylated (McConnell et al., 2018). In this study, three glycogen metabolism proteins, GPH1, GSN1, and TPS1, and one CMGC/MAPK/ERK protein kinase, PMK1, which are identified as redox modified, can be modified by phosphorylation (Franck et al., 2015). Moreover, GPH1 was reported to be N-glycosylated as well (Chen et al., 2020). Besides GPH1, there are other 14 proteins that can be modified by both redox and N-glycosylation, including the trehalase TRE1, chitin-binding protein CBP1, small GTPases Rac1, translocon channel protein MoSec62 and Sec5, and nine cell-wall-related proteins: MgCHS1, MgCHS4, MgCHS5, MgCHS6, Gel1, Gel2, Gel3, Gel4, and MoGls2 (Chen et al., 2020). Five proteins, including the cell wall biogenesis protein phosphatase SSD1, enoyl-CoA hydratase Ech1, heat shock proteins MoSsb1, isocitrate lyas ICL1, and the key enzymes in the glycolysis pathway, fructose-1,6-bisphosphatase FBP1, can be modified by both redox and succinylation (Wang et al., 2019). One protein, the heat shock proteins MoSsa1, can be modified by both redox and acetylation, with as many as 13 acetylation sites (Liang et al., 2018). The different modifications located on the same pathogenic crucial protein, suggesting a complicated mechanism of post-translational modification in pathogenicity regulation, may be possible by the interaction or crosstalk between different modifications.

The presence of thiol oxidation in the pathogenesis-related proteins indicated a role of redox modification in development, stress response, and pathogenicity of *M. oryzae* (**Figure 5**). However, further investigation is required for more details on the mechanism of the redox modification on these target proteins. The redox modification during each development stages, especially the infection hypha stages when deep interactions take place between the host and *M. oryzae*.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: The mass spectrometry proteomic datasets presented in this study can be found in online repository ProteomeXchange consortium via the PRIDE partner repository with an accession number of px-submission #470462.

AUTHOR CONTRIBUTIONS

XZ performed the experiments of sample preparation. X-LC designed the project. All authors performed the data processing and wrote the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grants 31871909 and 32072365) and the Open Research Fund of the State Key Laboratory of Hybrid Rice (Hunan Hybrid Rice Research Center) (2019KF04).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.648894/full#supplementary-material

Supplementary Table 1 | List of all thiol redox modified sites and proteins.

Supplementary Table 2 | Bioinformatics analyses data of the thiol redox modified proteins.

Supplementary Table 3 | Redox-modified pathogenesis-related proteins.

Supplementary Table 4 | Thiol redox-modified predicted effector proteins of *M. oryzae*.

REFERENCES

Badaruddin, M., Holcombe, L. J., Wilson, R. A., Wang, Z. Y., Kershaw, M. J., and Talbot, N. J. (2013). Glycogen metabolic genes are involved in trehalose-6-phosphate synthase-mediated regulation of pathogenicity by the rice blast fungus *Magnaporthe oryzae*. *PLoS Pathog.* 9:e1003604. doi: 10.1371/journal.ppat.1003604

Bae, Y. S., Oh, H., Rhee, S. G., and Yoo, Y. D. (2011). Regulation of reactive oxygen species generation in cell signaling. *Mol. Cells* 32, 491–509. doi: 10.1007/s10059-011-0276-3

Baraibar, M. A., and Friguet, B. (2013). Oxidative proteome modifications target specific cellular pathways during oxidative stress, cellular senescence and aging. *Exp. Gerontol.* 48, 620–625. doi: 10.1016/j.exger.2012.10.007

Beckerman, J. L., and Ebbole, D. J. (1996). MPG1, a gene encoding a fungal Hydrophobin of *Magnaporthe grisea*, is involved in surface recognition. *Mol. Plant Microb. Interact.* 9, 450–456. doi: 10.1094/mpmi-9-0450

Berepiki, A., Lichius, A., and Read, N. D. (2011). Actin organization and dynamics in filamentous fungi. *Nat. Rev. Microbiol.* 9, 876–887. doi: 10.1038/nrmicro2666

Bykova, N. V., Hoehn, B., Rampitsch, C., Banks, T., Stebbing, J. A., Fan, T., et al. (2011). Redox-sensitive proteome and antioxidant strategies in wheat seed dormancy control. *Proteomics* 11, 865–882. doi: 10.1002/pmic.200900810

Cai, X., Zhang, X., Li, X., Liu, M., Liu, X., Wang, X., et al. (2017). The atypical guanylate kinase MoGuk2 plays important roles in asexual/sexual development, conidial septation, and pathogenicity in the rice blast fungus. *Front. Microbiol.* 8:2467. doi: 10.3389/fmicb.2017.02467

Chen, J., Zheng, W., Zheng, S., Zhang, D., Sang, W., Chen, X., et al. (2008). Rac1 is required for pathogenicity and Chm1-dependent conidiogenesis in rice fungal pathogen *Magnaporthe grisea*. *PLoS Pathog.* 4:e1000202. doi: 10.1371/journal.ppat.1000202

Chen, X. L., Liu, C., Tang, B., Ren, Z., Wang, G. L., and Liu, W. (2020). Quantitative proteomics analysis reveals important roles of N-glycosylation on ER quality control system for development and pathogenesis in *Magnaporthe oryzae*. *PLoS Pathog.* 16:e1008355. doi: 10.1371/journal.ppat.1008355

Chen, X. L., Shi, T., Yang, J., Shi, W., Gao, X., Chen, D., et al. (2014). N-glycosylation of effector proteins by an alpha-1,3-mannosyltransferase is required for the rice blast fungus to evade host innate immunity. *Plant Cell* 26, 1360–1376. doi: 10.1105/tpc.114.123588

Cheng, A., Grant, C. E., Noble, W. S., and Bailey, T. L. (2019). MoMo: discovery of statistically significant post-translational modification motifs. *Bioinformatics* 35, 2774–2782. doi: 10.1093/bioinformatics/bty1058

Corcoran, A., and Cotter, T. G. (2013). Redox regulation of protein kinases. *FEBS J.* 280, 1944–1965. doi: 10.1111/febs.12224

Dagdas, Y. F., Yoshino, K., Dagdas, G., Ryder, L. S., Bielska, E., Steinberg, G., et al. (2012). Septin-mediated plant cell invasion by the rice blast fungus, *Magnaporthe oryzae*. *Science* 336, 1590–1595. doi: 10.1126/science.1222934

Darnell, J. E. Jr., Kerr, I. M., and Stark, G. R. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* 264, 1415–1421. doi: 10.1126/science.8197455

D'Autreaux, B., and Toledano, M. B. (2007). ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nat. Rev. Mol. Cell Biol.* 8, 813–824. doi: 10.1038/nrm2256

Dean, R., Van Kan, J. A., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., et al. (2012). The Top 10 fungal pathogens in molecular plant pathology. *Mol. Plant Pathol.* 13, 414–430. doi: 10.1111/j.1364-3703.2011.00783.x

Deng, Y. Z., and Naqvi, N. I. (2010). A vacuolar glucoamylase, Sga1, participates in glycogen autophagy for proper asexual differentiation in *Magnaporthe oryzae*. *Autophagy* 6, 455–461. doi: 10.4161/auto.6.4.11736

Deng, Y. Z., Ramos-Pamplona, M., and Naqvi, N. I. (2009). Autophagy-assisted glycogen catabolism regulates asexual differentiation in *Magnaporthe oryzae*. *Autophagy* 5, 33–43. doi: 10.4161/auto.5.1.7175

Fass, E., Amar, N., and Elazar, Z. (2007). Identification of essential residues for the C-terminal cleavage of the mammalian LC3: a lesson from yeast Atg8. *Autophagy* 3, 48–50. doi: 10.4161/auto.3417

Foster, A. J., Jenkinson, J. M., and Talbot, N. J. (2003). Trehalose synthesis and metabolism are required at different stages of plant infection by *Magnaporthe grisea*. *EMBO J.* 22, 225–235. doi: 10.1093/emboj/cdg018

Franck, W. L., Gokce, E., Randall, S. M., Oh, Y., Eyre, A., Muddiman, D. C., et al. (2015). Phosphoproteome analysis links protein phosphorylation to cellular remodeling and metabolic adaptation during *Magnaporthe oryzae* appressorium development. *J. Proteome Res.* 14, 2408–2424. doi: 10.1021/pr501064q

Furuya, N., Yu, J., Byfield, M., Pattingre, S., and Levine, B. (2005). The evolutionarily conserved domain of Beclin 1 is required for Vps34 binding, autophagy and tumor suppressor function. *Autophagy* 1, 46–52. doi: 10.4161/auto.1.1.1542

Giorgio, M., Migliaccio, E., Orsini, F., Paolucci, D., Moroni, M., Contursi, C., et al. (2005). Electron transfer between cytochrome c and p66Shc generates reactive oxygen species that trigger mitochondrial apoptosis. *Cell* 122, 221–233. doi: 10.1016/j.cell.2005.05.011

Go, Y. M., and Jones, D. P. (2013). The redox proteome. *J. Biol. Chem.* 288, 26512–26520.

Guo, J., Nguyen, A. Y., Dai, Z., Su, D., Gaffrey, M. J., Moore, R. J., et al. (2014). Proteome-wide light/dark modulation of thiol oxidation in cyanobacteria revealed by quantitative site-specific redox proteomics. *Mol. Cell Proteom.* 13, 3270–3285. doi: 10.1074/mcp.m114.041160

Guo, M., Gao, F., Zhu, X., Nie, X., Pan, Y., and Gao, Z. (2015). MoGrr1, a novel F-box protein, is involved in conidiogenesis and cell wall integrity and is critical for the full virulence of *Magnaporthe oryzae*. *Appl. Microbiol. Biotechnol.* 99, 8075–8088. doi: 10.1007/s00253-015-6820-x

Hancock, J. T. (2009). The role of redox mechanisms in cell signalling. *Mol. Biotechnol.* 43, 162–166. doi: 10.1007/s12033-009-9189-1

He, M., Xu, Y., Chen, J., Luo, Y., Lv, Y., Su, J., et al. (2018). MoSnt2-dependent deacetylation of histone H3 mediates MoTor-dependent autophagy and plant infection by the rice blast fungus *Magnaporthe oryzae*. *Autophagy* 14, 1543–1561. doi: 10.1080/15548627.2018.1458171

He, Y., Deng, Y. Z., and Naqvi, N. I. (2013). Atg24-assisted mitophagy in the foot cells is necessary for proper asexual differentiation in *Magnaporthe oryzae*. *Autophagy* 9, 1818–1827. doi: 10.4161/auto.26057

Heppner, D. E., Dustin, C. M., Liao, C., Hristova, M., Veith, C., Little, A. C., et al. (2018). Direct cysteine sulfenylation drives activation of the Src kinase. *Nat. Commun.* 9:4522.

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., et al. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 35, W585–W587.

Huntley, R. P., Sawford, T., Mutowo-Meullenet, P., Shypitsyna, A., Bonilla, C., Martin, M. J., et al. (2015). The GOA database: gene Ontology annotation updates for 2015. *Nucleic Acids Res.* 43, D1057–D1063.

Jarosch, E., Tuller, G., Daum, G., Waldherr, M., Voskova, A., and Schweyen, R. J. (1996). Mrs5p, an essential protein of the mitochondrial intermembrane space, affects protein import into yeast mitochondria. *J. Biol. Chem.* 271, 17219–17225. doi: 10.1074/jbc.271.29.17219

Jeon, J., Goh, J., Yoo, S., Chi, M. H., Choi, J., Rho, H. S., et al. (2008). A putative MAP kinase kinase kinase, MCK1, is required for cell wall integrity and pathogenicity of the rice blast fungus, *Magnaporthe oryzae*. *Mol. Plant Microb. Interact.* 21, 525–534. doi: 10.1094/mpmi-21-5-0525

Jeong, J. S., Mitchell, T. K., and Dean, R. A. (2007). The *Magnaporthe grisea* snodprot1 homolog, MSP1, is required for virulence. *FEMS Microbiol. Lett.* 273, 157–165. doi: 10.1111/j.1574-6968.2007.00796.x

Kanehisa, M., and Sato, Y. (2020). KEGG Mapper for inferring cellular functions from protein sequences. *Protein Sci.* 29, 28–35. doi: 10.1002/pro.3711

Kou, Y., Qiu, J., and Tao, Z. (2019). Every coin has two sides: reactive oxygen species during rice-*Magnaporthe oryzae* interaction. *Int. J. Mol. Sci.* 20:1191. doi: 10.3390/ijms20051191

Kumsta, C., Thamsen, M., and Jakob, U. (2011). Effects of oxidative stress on behavior, physiology, and the redox thiol proteome of *Caenorhabditis elegans*. *Antioxid. Redox Signal.* 14, 1023–1037. doi: 10.1089/ars.2010.3203

Li, Y., Yue, X., Que, Y., Yan, X., Ma, Z., Talbot, N. J., et al. (2014). Characterisation of four LIM protein-encoding genes involved in infection-related development and pathogenicity by the rice blast fungus *Magnaporthe oryzae*. *PLoS One* 9:e88246. doi: 10.1371/journal.pone.0088246

Liang, M., Zhang, S., Dong, L., Kou, Y., Lin, C., Dai, W., et al. (2018). Label-free quantitative proteomics of lysine acetylome identifies substrates of gcn5 in *Magnaporthe oryzae* autophagy and epigenetic regulation. *mSystems* 3:e0270-18.

Lim, Y. J., Kim, K. T., and Lee, Y. H. (2018). SUMOylation is required for fungal development and pathogenicity in the rice blast fungus *Magnaporthe oryzae*. *Mol. Plant Pathol.* 19, 2134–2148. doi: 10.1111/mpp.12687

Liu, P., Zhang, H., Wang, H., and Xia, Y. (2014). Identification of redox-sensitive cysteines in the *Arabidopsis* proteome using OxiTRAQ, a quantitative redox proteomics method. *Proteomics* 14, 750–762. doi: 10.1002/pmic.201300307

Liu, S., and Dean, R. A. (1997). G protein alpha subunit genes control growth, development, and pathogenicity of *Magnaporthe grisea*. *Mol. Plant Microb. Interact.* 10, 1075–1086. doi: 10.1094/mpmi.1997.10.9.1075

Liu, X. H., Chen, S. M., Gao, H. M., Ning, G. A., Shi, H. B., Wang, Y., et al. (2015). The small GTPase MoYpt7 is required for membrane fusion in autophagy and pathogenicity of *Magnaporthe oryzae*. *Environ. Microbiol.* 17, 4495–4510. doi: 10.1111/1462-2920.12903

Liu, X. H., Lu, J. P., Dong, B., Gu, Y., and Lin, F. C. (2010). Disruption of MoCMK1, encoding a putative calcium/calmodulin-dependent kinase, in *Magnaporthe oryzae*. *Microbiol. Res.* 165, 402–410. doi: 10.1016/j.micres.2009.08.007

Liu, X. H., Lu, J. P., Zhang, L., Dong, B., Min, H., and Lin, F. C. (2007). Involvement of a *Magnaporthe grisea* serine/threonine kinase gene, MgATG1, in appressorium turgor and pathogenesis. *Eukaryot Cell* 6, 997–1005. doi: 10.1128/ec.00011-07

Maeda, Y., Oku, M., and Sakai, Y. (2015). A defect of the vacuolar putative lipase Atg15 accelerates degradation of lipid droplets through lipolysis. *Autophagy* 11, 1247–1258. doi: 10.1080/15548627.2015.1056969

McConnell, E. W., Werth, E. G., and Hicks, L. M. (2018). The phosphorylated redox proteome of *Chlamydomonas reinhardtii*: Revealing novel means for regulation of protein structure and function. *Redox Biol.* 17, 35–46. doi: 10.1016/j.redox.2018.04.003

McDonagh, B., Dominguez-Martin, M. A., Gomez-Baena, G., Lopez-Lozano, A., Diez, J., Barcena, J. A., et al. (2012). Nitrogen starvation induces extensive changes in the redox proteome of *Prochlorococcus* sp. strain SS120. *Environ. Microbiol. Rep.* 4, 257–267. doi: 10.1111/j.1758-2229.2012.00329.x

Mogga, V., Delventhal, R., Weidenbach, D., Langer, S., Bertram, P. M., Andresen, K., et al. (2016). *Magnaporthe oryzae* effectors MoHEG13 and MoHEG16 interfere with host infection and MoHEG13 counteracts cell death caused by *Magnaporthe*-NLPs in tobacco. *Plant Cell Rep.* 35, 1169–1185. doi: 10.1007/s00299-016-1943-9

Moseley, J. B., and Goode, B. L. (2006). The yeast actin cytoskeleton: from cellular function to biochemical mechanism. *Microbiol. Mol. Biol. Rev.* 70, 605–645. doi: 10.1128/mmbr.00013-06

Murray, C. I., and Van Eyk, J. E. (2012). Chasing cysteine oxidative modifications: proteomic tools for characterizing cysteine redox status. *Circ. Cardiovasc. Genet.* 5:591. doi: 10.1161/circgenetics.111.961425

Nakatogawa, H., Ichimura, Y., and Ohsumi, Y. (2007). Atg8, a ubiquitin-like protein required for autophagosome formation, mediates membrane tethering and hemifusion. *Cell* 130, 165–178. doi: 10.1016/j.cell.2007.05.021

Nie, H., Zhang, L., Zhuang, H., Yang, X., Qiu, D., and Zeng, H. (2019). Secreted protein MoHrip2 is required for full virulence of *Magnaporthe oryzae* and modulation of rice immunity. *Appl. Microbiol. Biotechnol.* 103, 6153–6167. doi: 10.1007/s00253-019-09937-2

Olsen, J. V., and Mann, M. (2013). Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol. Cell Proteom.* 12, 3444–3452. doi: 10.1074/mcp.o113.034181

Panagabko, C., Morley, S., Hernandez, M., Cassolato, P., Gordon, H., Parsons, R., et al. (2003). Ligand specificity in the CRAL-TRIO protein family. *Biochemistry* 42, 6467–6474. doi: 10.1021/bi034086v

Paschen, S. A., Rothbauer, U., Kaldi, K., Bauer, M. F., Neupert, W., and Brunner, M. (2000). The role of the TIM8-13 complex in the import of Tim23 into mitochondria. *EMBO J.* 19, 6392–6400. doi: 10.1093/emboj/19.23.6392

Paulsen, C. E., and Carroll, K. S. (2010). Orchestrating redox signaling networks through regulatory cysteine switches. *ACS Chem. Biol.* 5, 47–62. doi: 10.1021/cb900258z

Prakash, C., Manjrekar, J., and Chattoo, B. B. (2016). Skp1, a component of E3 ubiquitin ligase, is necessary for growth, sporulation, development and pathogenicity in rice blast fungus (*Magnaporthe oryzae*). *Mol. Plant Pathol.* 17, 903–919. doi: 10.1111/mpp.12336

Qian, B., Liu, X., Jia, J., Cai, Y., Chen, C., Zhang, H., et al. (2018). MoPpe1 partners with MoSap1 to mediate TOR and cell wall integrity signalling in growth and pathogenicity of the rice blast fungus *Magnaporthe oryzae*. *Environ. Microbiol.* 20, 3964–3979. doi: 10.1111/1462-2920.14421

Qiu, X. B., Shao, Y. M., Miao, S., and Wang, L. (2006). The diversity of the DnaJ/Hsp40 family, the crucial partners for Hsp70 chaperones. *Cell Mol. Life Sci.* 63, 2560–2570. doi: 10.1007/s00018-006-6192-6

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., et al. (2005). InterProScan: protein domains identifier. *Nucleic Acids Res.* 33, W116–W120.

Rigoulet, M., Yoboue, E. D., and Devin, A. (2011). Mitochondrial ROS generation and its regulation: mechanisms involved in H_2O_2 signaling. *Antioxid. Redox Signal.* 14, 459–468. doi: 10.1089/ars.2010.3363

Rosenwasser, S., Graff van Creveld, S., Schatz, D., Malitsky, S., Tzfadia, O., Aharoni, A., et al. (2014). Mapping the diatom redox-sensitive proteome provides insight into response to nitrogen stress in the marine environment. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2740–2745. doi: 10.1073/pnas.1319773111

Shi, H. B., Chen, G. Q., Chen, Y. P., Dong, B., Lu, J. P., Liu, X. H., et al. (2016). MoRad6-mediated ubiquitination pathways are essential for development and pathogenicity in *Magnaporthe oryzae*. *Environ. Microbiol.* 18, 4170–4187. doi: 10.1111/1462-2920.13515

Sies, H. (2014). Role of metabolic H_2O_2 generation: redox signaling and oxidative stress. *J. Biol. Chem.* 289, 8735–8741.

Sies, H. (2017). Hydrogen peroxide as a central redox signaling molecule in physiological oxidative stress: oxidative eustress. *Redox Biol.* 11, 613–619. doi: 10.1016/j.redox.2016.12.035

Sievers, S., Dittmann, S., Jordt, T., Otto, A., Hochgrafe, F., and Riedel, K. (2018). Comprehensive redox profiling of the thiol proteome of *Clostridium difficile*. *Mol. Cell Proteom.* 17, 1035–1046. doi: 10.1074/mcp.tir118.000671

Silacci, P., Mazzolai, L., Gauci, C., Stergiopulos, N., Yin, H. L., and Hayoz, D. (2004). Gelsolin superfamily proteins: key regulators of cellular functions. *Cell Mol. Life Sci.* 61, 2614–2623. doi: 10.1007/s00018-004-4225-6

Su, Z., Burchfield, J. G., Yang, P., Humphrey, S. J., Yang, G., Francis, D., et al. (2019). Global redox proteome and phosphoproteome analysis reveals redox switch in Akt. *Nat. Commun.* 10:5486.

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613.

Talbot, N. J., Ebbole, D. J., and Hamer, J. E. (1993). Identification and characterization of MPG1, a gene involved in pathogenicity from the rice blast fungus *Magnaporthe grisea*. *Plant Cell* 5, 1575–1590. doi: 10.2307/3869740

Talbot, N. J., Kershaw, M. J., Wakley, G. E., De Vries, O., Wessels, J., and Hamer, J. E. (1996). MPG1 encodes a fungal hydrophobin involved in surface interactions during infection-related development of *Magnaporthe grisea*. *Plant Cell* 8, 985–999. doi: 10.1105/tpc.8.6.985

Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36. doi: 10.1093/nar/28.1.33

Thangima Zannat, M., Bhattacharjee, R. B., and Bag, J. (2011). In the absence of cellular poly (A) binding protein, the glycolytic enzyme GAPDH translocated to the cell nucleus and activated the GAPDH mediated apoptotic pathway by enhancing acetylation and serine 46 phosphorylation of p53. *Biochem. Biophys. Res. Commun.* 409, 171–176. doi: 10.1016/j.bbrc.2011.04.094

Thines, E., Weber, R. W., and Talbot, N. J. (2000). MAP kinase and protein kinase A-dependent mobilization of triacylglycerol and glycogen during appressorium turgor generation by *Magnaporthe grisea*. *Plant Cell* 12, 1703–1718. doi: 10.2307/3871184

Tobiume, K., Matsuzawa, A., Takahashi, T., Nishitoh, H., Morita, K., Takeda, K., et al. (2001). ASK1 is required for sustained activations of JNK/p38 MAP kinases and apoptosis. *EMBO Rep.* 2, 222–228. doi: 10.1093/embo-reports/kve046

Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* 11, 2301–2319. doi: 10.1038/nprot.2016.136

Veneault-Fourrey, C., Barooah, M., Egan, M., Wakley, G., and Talbot, N. J. (2006). Autophagic fungal cell death is necessary for infection by the rice blast fungus. *Science* 312, 580–583. doi: 10.1126/science.1124550

Wang, J., Li, L., Chai, R., Zhang, Z., Qiu, H., Mao, X., et al. (2019). Succinyl-proteome profiling of *Pyricularia oryzae*, a devastating phytopathogenic fungus that causes rice blast disease. *Sci. Rep*. 9:3490.

Wang, Z., Zhang, H., Liu, C., Xing, J., and Chen, X. L. (2018). A deubiquitinating enzyme Ubp14 is required for development, stress response, nutrient utilization, and pathogenesis of *Magnaporthe oryzae*. *Front. Microbiol.* 9:769. doi: 10.3389/fmicb.2018.00769

Wilson, R. A., and Talbot, N. J. (2009). Under pressure: investigating the biology of plant infection by *Magnaporthe oryzae*. *Nat. Rev. Microbiol.* 7, 185–195. doi: 10.1038/nrmicro2032

Winterbourn, C. C., and Hampton, M. B. (2008). Thiol chemistry and specificity in redox signaling. *Free Radic. Biol. Med.* 45, 549–561. doi: 10.1016/j.freeradbiomed.2008.05.004

Xie, K., Bunse, C., Marcus, K., and Leichert, L. I. (2019). Quantifying changes in the bacterial thiol redox proteome during host-pathogen interaction. *Redox Biol.* 21:101087. doi: 10.1016/j.redox.2018.101087

Xu, J. R., and Hamer, J. E. (1996). MAP kinase and cAMP signaling regulate infection structure formation and pathogenic growth in the rice blast fungus *Magnaporthe grisea*. *Genes Dev.* 10, 2696–2706. doi: 10.1101/gad.10.21.2696

Xu, J. R., Staiger, C. J., and Hamer, J. E. (1998). Inactivation of the mitogen-activated protein kinase Mps1 from the rice blast fungus prevents penetration of host cells but allows activation of plant defense responses. *Proc. Natl. Acad. Sci. U.S.A.* 95, 12713–12718. doi: 10.1073/pnas.95.21.12713

Yang, C. D., Dang, X., Zheng, H. W., Chen, X. F., Lin, X. L., Zhang, D. M., et al. (2017). Two Rab5 homologs are essential for the development and pathogenicity of the rice blast fungus *Magnaporthe oryzae*. *Front. Plant Sci.* 8:620. doi: 10.3389/fpls.2017.00620

Yen, W. L., Legakis, J. E., Nair, U., and Klionsky, D. J. (2007). Atg27 is required for autophagy-dependent cycling of Atg9. *Mol. Biol. Cell* 18, 581–593. doi: 10.1091/mbc.e06-07-0612

Yi, M., Park, J. H., Ahn, J. H., and Lee, Y. H. (2008). MoSNF1 regulates sporulation and pathogenicity in the rice blast fungus *Magnaporthe oryzae*. *Fungal Genet. Biol.* 45, 1172–1181. doi: 10.1016/j.fgb.2008.05.003

Yin, Z., Balmant, K., Geng, S., Zhu, N., Zhang, T., Dufresne, C., et al. (2017). Bicarbonate induced redox proteome changes in *Arabidopsis* suspension cells. *Front. Plant Sci.* 8:58. doi: 10.3389/fpls.2017.00058

Yin, Z., Tang, W., Wang, J., Liu, X., Yang, L., Gao, C., et al. (2016). Phosphodiesterase MoPdeH targets MoMck1 of the conserved mitogen-activated protein (MAP) kinase signalling pathway to regulate cell wall integrity in rice blast fungus *Magnaporthe oryzae*. *Mol. Plant Pathol.* 17, 654–668. doi: 10.1111/mpp.12317

Zhang, H., Liu, K., Zhang, X., Song, W., Zhao, Q., Dong, Y., et al. (2010). A two-component histidine kinase, MoSLN1, is required for cell wall integrity and pathogenicity of the rice blast fungus, *Magnaporthe oryzae*. *Curr. Genet.* 56, 517–528. doi: 10.1007/s00294-010-0319-x

Zhang, S., Liu, X., Li, L., Yu, R., He, J., Zhang, H., et al. (2017). The ArfGAP protein MoGlo3 regulates the development and pathogenicity of *Magnaporthe oryzae*. *Environ. Microbiol.* 19, 3982–3996. doi: 10.1111/1462-2920.13798

Zhao, X., Kim, Y., Park, G., and Xu, J. R. (2005). A mitogen-activated protein kinase cascade regulating infection-related morphogenesis in *Magnaporthe grisea*. *Plant Cell* 17, 1317–1329. doi: 10.1105/tpc.104.029116

Zhao, Y., and Jensen, O. N. (2009). Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* 9, 4632–4641. doi: 10.1002/pmic.200900398

Zheng, W., Chen, J., Liu, W., Zheng, S., Zhou, J., Lu, G., et al. (2007). A Rho3 homolog is essential for appressorium development and pathogenicity of *Magnaporthe grisea*. *Eukaryot Cell* 6, 2240–2250. doi: 10.1128/ec.00104-07

Zheng, W., Chen, J. S., Zheng, S. Q., Liu, W. D., Chen, X., Zhou, J., et al. (2006). Relationship of Rho GTPase in *Magnaporthe griesea* revealed by gene expression analysis. *Sci. Agric. Sin*. 39, 2237–2242.

Zheng, W., Zhao, Z., Chen, J., Liu, W., Ke, H., Zhou, J., et al. (2009). A Cdc42 ortholog is required for penetration and virulence of *Magnaporthe grisea*. *Fungal Genet. Biol.* 46, 450–460. doi: 10.1016/j.fgb.2009.03.005

Zhou, T., Dagdas, Y. F., Zhu, X., Zheng, S., Chen, L., Cartwright, Z., et al. (2017). The glycogen synthase kinase MoGsk1, regulated by Mps1 MAP kinase, is required for fungal development and pathogenicity in *Magnaporthe oryzae*. *Sci. Rep.* 7:945.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 02 July 2021
doi: 10.3389/fmicb.2021.679936

Comparative Analysis of Host-Associated Variation in *Phytophthora cactorum*

Charlotte F. Nellist[1,2*†‡], Andrew D. Armitage[1,3†‡], Helen J. Bates[1,2†], Maria K. Sobczyk[1†§], Matteo Luberti[1†], Laura A. Lewis[1§] and Richard J. Harrison[1,2*†]

[1] NIAB EMR, East Malling, United Kingdom, [2] NIAB, Cambridge, United Kingdom, [3] National Resources Institute, University of Greenwich, Chatham, United Kingdom

Phytophthora cactorum is often described as a generalist pathogen, with isolates causing disease in a range of plant species. It is the causative agent of two diseases in the cultivated strawberry, crown rot (CR; causing whole plant collapse) and leather rot (LR; affecting the fruit). In the cultivated apple, *P. cactorum* causes girdling bark rots on the scion (collar rot) and rootstock (crown rot), as well as necrosis of the fine root system (root rot) and fruit rots. We investigated evidence for host specialisation within *P. cactorum* through comparative genomic analysis of 18 isolates. Whole genome phylogenetic analysis provided genomic support for discrete lineages within *P. cactorum*, with well-supported non-recombining clades for strawberry CR and apple infecting isolates specialised to strawberry crowns and apple tissue. Isolates of strawberry CR are genetically similar globally, while there is more diversity in apple-infecting isolates. We sought to identify the genetic basis of host specialisation, demonstrating gain and loss of effector complements within the *P. cactorum* phylogeny, representing putative determinants of host boundaries. Transcriptomic analysis highlighted that those effectors found to be specific to a single host or expanded in the strawberry lineage are amongst those most highly expressed during infection of strawberry and give a wider insight into the key effectors active during strawberry infection. Many effectors that had homologues in other *Phytophthoras* that have been characterised as avirulence genes were present but not expressed in our tested isolate. Our results highlight several RxLR-containing effectors that warrant further investigation to determine whether they are indeed virulence factors and host-specificity determinants for strawberry and apple. Furthermore, additional work is required to determine whether these effectors are suitable targets to focus attention on for future resistance breeding efforts.

Keywords: crown rot, oomycete, phylogenomics, effectors, cryptic species

OPEN ACCESS

Edited by:
Silvia Laura Toffolatti, University of Milan, Italy

Reviewed by:
Natalia Peres, University of Florida, United States
Franck Panabieres, INRA Centre Provence-Alpes-Côte d'Azur, France
Seonghee Lee, University of Florida, United States
Tomasz Oszako, Forest Research Institute (IBL), Poland

***Correspondence:**
Charlotte F. Nellist
Charlotte.Nellist@niab.com
Richard J. Harrison
Richard.Harrison@niab.com

†ORCID:
Charlotte F. Nellist
orcid.org/0000-0002-7453-3710
Andrew D. Armitage
orcid.org/0000-0002-0610-763X
Helen J. Bates
orcid.org/0000-0002-7978-9213
Maria K. Sobczyk
orcid.org/0000-0003-3329-612X
Matteo Luberti
orcid.org/0000-0002-3667-2920
Richard J. Harrison
orcid.org/0000-0002-3307-3519

‡*These authors have contributed equally to this work*

§***Present address:***
Maria K. Sobczyk, MRC-IEU, Bristol Medical School, University of Bristol, Oakfield House, Bristol, United Kingdom
Laura A. Lewis, The Rhodes Trust, Rhodes House, Oxford, United Kingdom

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *12 March 2021*
Accepted: *21 May 2021*
Published: *02 July 2021*

INTRODUCTION

The *Phytophthora* genus serves as a model system for studying evolution of pathogenicity and resistance in plant pathosystems. Over 150 species have been named in the genus, with many pathogenic on plants (Yang et al., 2017). *Phytophthora* spp. are extremely effective plant pathogens that are able to disperse and infect hosts via the asexual, motile stage of their life cycle, zoospores, as well as able to resist and survive for many years in unfavourable conditions as thick-walled

sexual oospores. Many *Phytophthora* spp. are specialised and only able to colonise one or a few host plants, for example *Phytophthora fragariae* which is thought to only colonise strawberry. Some species, such as *Phytophthora cactorum* (Lebert and Cohn) Schroet, are traditionally considered to be generalists and are able to cause disease on a broad range of plant species, including herbaceous and woody plants (Erwin and Ribeiro, 1996). Two examples of these hosts in the *Rosaceae* family, are the herbaceous cultivated strawberry (*Fragaria* × *ananassa*) and woody cultivated apple (*Malus* × *domestica*).

Phytophthora cactorum is the causative agent of two diseases in the cultivated strawberry, crown rot (CR; causing whole plant collapse (Deutschmann, 1954); and leather rot (LR; affecting the fruit (Rose, 1924). Both are considered major diseases of strawberry in temperate regions, with crop losses of up to 40 and 30% reported, respectively, for each disease (Ellis and Grove, 1983; Stensvand et al., 1999). The majority of commercial strawberries grown in the UK are grown under polytunnels or in glasshouses, on tabletops using soilless substrate, such as coir (coconut husk) (Boyer et al., 2016). *P. cactorum* is a particular problem in this production system due to ease of spread through the irrigation system via the motile asexual life-stage of *Phytophthora*, zoospores. Introduction into the irrigation system through asymptomatic infections in planting material represents the biggest risk to growers. Nursery propagation of strawberry is still based in the field, where the presence of resident pathogen inoculum in the soil along with latent infection in plants represent a severe threat to production. A study in 2018 of UK strawberry planting material (runners), commissioned by the Agriculture and Horticulture Development Board (AHDB), reported incidences of *P. cactorum* as high as 30% and great variation observed between batches of plants tested, but on average an incidence of 8–10% was recorded (Wedgwood et al., 2020).

In the cultivated apple, *P. cactorum* causes girdling bark rots on the scion (collar rot) and rootstock (crown rot), as well as necrosis of the fine root system (root rot) and fruit rots (Harris, 1991). *P. cactorum* affects nearly all apple growing regions of the world, debilitating the trees and leading to reduced fruit yield and eventual tree death (Alexander and Stewart, 2001). Due to the high costs associated with orchard establishment, losses due to *P. cactorum* can result in significant economic losses. As *P. cactorum* is homothallic (self-fertile), it produces sexual oospores which can survive for long periods in the soil, growing material and plant material, contributing to its importance as a worldwide threat to apple production.

Phytophthora cactorum diverged from other Clade 1 *Phytophthora* spp., *P. infestans* and *P. parasitica*, an estimated 221.4 Ma (138.6–342.4 million years ago; Yang et al., 2018) which is some 100 My (million years) earlier than the divergence of the Dryadoideae, Ammygdaloieae, and Rosioideae and some 150–170 My before the emergence of the Fragariae (Zhang et al., 2017). *P. cactorum* specifically belongs to subclade 1a, along with *P. idaei* (Yang et al., 2017), a sister taxon and pathogen of another important *Rosaceae* crop, raspberry (*Rubus idaeus*). Despite such a broad host range being described for *P. cactorum*, host specialisation has been documented to particular plant species, including strawberry and apple (Seemüller and Schmidle, 1979). *P. cactorum* isolates originating from strawberry crowns were found to be less pathogenic on apple bark tissue than isolates originating from strawberry fruit or apple and vice versa, with apple and strawberry fruit isolates being less pathogenic on strawberry crowns. It was found that all pathotypes were able to cause disease in strawberry fruit (Seemüller and Schmidle, 1979). *P. cactorum* host specialisation has also been reported in other hosts, such as silver birch, peach and almond (Hantula et al., 1997, 2000; Lilja et al., 1998; Thomidis, 2003; Bhat et al., 2006). The specific genetic components responsible for host specialisation in *P. cactorum* are not known. Although, studies in both filamentous and bacterial pathogen systems support the model of effector repertoire being one of the key determinants in pathogen host range and non-host resistance (Schulze-Lefert and Panstruga, 2011; Stam et al., 2014).

Genomic resources have recently become available for *P. cactorum*, with genomes released for isolates from beech, *Fagus sylvatica* (Grenville-Briggs et al., 2017), Chinese ginseng, *Panax notoginseng* (Yang et al., 2018) and strawberry, *F. x ananassa* (Armitage et al., 2018). *Phytophthora* spp. carry two major families of cytoplasmic effectors that are translocated into the host cell, RxLR and Crinklers, both have been characterised by specific motifs and consist of a rapidly evolving effector/modulating domain. The RxLR family of effectors are characterised by an N-terminal signal peptide, followed by an RxLR-s/dEER motif (Arginine, any amino-acid, Leucine, Arginine often followed by Serine/Aspartate, Glutamate, Glutamate, Arginine) and a variable C-terminal domain that may contain WY domain repeats (Wawra et al., 2012; Win et al., 2012). These effectors are renowned for suppressing host defence mechanisms through the manipulation of various aspects of plant defence (Anderson et al., 2015). The second family of cytoplasmic effectors are Crinklers (CRNs), named because of the leaf crinkling effect observed when expressed in host plants (Torto et al., 2003). They are characterised by an N-terminal LFLAK domain followed by a DWL domain, with a DI domain sometimes present between these two domains (Haas et al., 2009; Stam et al., 2013). These conserved domains are followed by variable C-terminal domains.

In addition to cytoplasmic effectors, *Phytophthora* spp. also deploy an arsenal of apoplastic effector proteins during infection, including a large number of hydrolytic enzymes, such as cutinases, glycoside hydrolases (GHs), pectinases, and proteases, which promote their infection and manipulation of the plant immune system (Armitage et al., 2018; Wang and Wang, 2018a,b). They also encode members of extracellular phytotoxin families such as the conserved necrosis-inducing proteins (NLPs) and small cysteine-rich (SCR) proteins, for such as PcF (*P. cactorum* factor) (Orsomando et al., 2001) and INF1 (Kamoun et al., 1997).

Here, we further explore host specialisation and the basis of pathogenicity in *P. cactorum* by investigating multiple isolates collected from symptomatic strawberry crowns and fruit, as well as isolates from symptomatic apple tissue. We demonstrate that there are distinct lineages within *P. cactorum* showing adaptation to strawberry crowns and apple tissue. We show that these

lineages are associated with unique effector complements and that these differential genes are highly expressed during plant infection. Taken together, this work elucidates key lineage specific effector genes playing roles in specialisation to strawberry and apple in *P. cactorum*.

MATERIALS AND METHODS

Phytophthora Isolates Investigated in This Study

Eighteen *P. cactorum* isolates and three *P. idaei* isolates (detailed in **Table 1**) were investigated in this study. Of the 18 *P. cactorum* isolates, 13 were isolated from strawberry crown tissue exhibiting crown rot symptoms, two from strawberry fruit exhibiting leather rot symptoms and three isolated from symptomatic apple bark. One of the crown rot isolates, 10300, has been previously published by Armitage et al. (2018). The three isolates of *P. idaei* were isolated from infected raspberry material. All isolates were revived and maintained on V8 agar at 20°C in the dark.

Whole Genome Sequencing and Assembly

Mycelia of *P. cactorum* and *P. idaei* were grown in clarified V8-juice broth, similar to Wilcox et al. (1993); comprised of 100 mL V8-juice (Arnotts Biscuits Limited), 1.4 g calcium carbonate ($CaCO_3$; Sigma Aldrich) and 100 mL dH_2O, which were centrifuged at 2,500 × *g* for 15 min, the supernatant was decanted and made up to 1,600 mL with dH_2O; 200 mL aliquots were dispensed into 250 mL flasks and autoclaved for 20 min at 120°C. Five mycelial plugs per isolate were added to each flask and were grown at 20°C under lab light/dark cycle for 10 days in a shaker incubator set to 200 rpm (revolutions per minute). The mycelia were washed in sterile dH_2O, vacuum filtered and freeze dried overnight.

The GenElute Plant Genomic DNA Kit (Sigma) was used to extract gDNA from *P. idaei* isolate, SCRP371. gDNA was sonicated in a water bath and size selected, ∼500 bp, on an agarose gel and extracted. An Illumina library was constructed using the TruSeq LT Kit (FC-121-2001) and was sequenced using Illumina MiSeq v3 2x 300 bp PE Reagent Kit. For all remaining isolates gDNA was extracted using the Macherey-Nagel Nucleospin Plant II Kit (Fisher Scientific). gDNA was sheared using the Covaris M220 with microTUBE-50 (Covaris) and size selected, 450−600 bp, using a Blue Pippin (Sage Science). Illumina libraries were constructed using a PCR-free method using NEBNext End Repair (E6050S), NEBNext dA-tailing (E6053S) and Blunt T/A ligase (M0367S) New England Biolabs modules. Library insert sizes were 400–600 bp and were sequenced using Illumina Miseq v2 2x 250 bp paired-end (PE; MS-102-2003) or v3 2x 300 bp PE (MS-102-3003) Reagent Kits.

A single *P. cactorum* isolate (P414), from a symptomatic crown of strawberry, was selected for additional PacBio sequencing. gDNA extraction was performed using the Genomic-tip DNA 100/G Kit (Qiagen), following the Tissue Sample method. A minimum of 20 μg of gDNA at approximately 100 ng/μL concentration, with a 260/280 ratio of 1.88 and a 260/230 ratio of 2.26, and a minimum molecular weight of 40 kb was sent to The Earlham Institute, UK. The large insert library was prepared by The Earlham Institute according to manufacturer specifications and sequenced to achieve approximately 87 times coverage on a PacBio RSII platform, using P6-C4 chemistry.

A long-read *de novo* assembly was generated for isolate P414 by first performing read correction and trimming using Canu v1.6 (Koren et al., 2017), before assembling with SMARTdenovo (February 26, 2017 github commit). Errors in this SMARTdenovo assembly were polished through five iterations of Pilon v1.17 (Walker et al., 2014), using the "diploid" flag and trimmed Illumina reads. Illumina reads were trimmed to remove low quality bases and Illumina adapters with fastq-mcf v1.04.676 (Aronesty, 2013).

De novo assembly of MiSeq data for the remaining 19 genomes was performed using SPAdes v.3.11.0 (Bankevich et al., 2012). Assembly statistics were collected for all assemblies using QUAST v3.0 (Gurevich et al., 2013). Completeness of the *Phytophthora* genome assemblies was assessed by analysis of conserved Benchmarking Universal Single-Copy Orthologue (BUSCO, v3; Simão et al., 2015; Waterhouse et al., 2017) genes using the Alveolata-Stramenopiles dataset. DeconSeq was run on all assemblies to remove any potential bacterial contaminants with homology to databases of all "complete" *Bacillus* or *Paenobacillus* genomes as downloaded from NCBI (Schmieder and Edwards, 2011). A database of *Phytophthora* contigs was also made and contigs that showed homology to both bacterial and *Phytophthora* databases were retained. Assemblies were edited in accordance with results from the NCBI contamination screen (run as part of submission to GenBank in December 2017) with contigs split, trimmed or excluded as required. RepeatModeler, RepeatMasker and transposonPSI were used to identify repetitive and low complexity regions (http://www.repeatmasker.org, http://transposonpsi.sourceforge.net).

Gene and Open Reading Frame Prediction and Functional Annotation

Gene prediction was performed following Armitage et al. (2018) and detailed is in **Supplementary Materials and Methods**. Gene models were also augmented with further effector candidates from open reading frames (ORFs) using the methods previously described in Armitage et al. (2018).

Functional annotation of gene models was performed as described previously in Armitage et al. (2018), further details can be found in **Supplementary Materials and Methods**. A publicly available *P. cactorum* genome, LV007, isolated from European Beech (*Fagus sylvatica*) was downloaded from GenBank (PRJNA380728; Grenville-Briggs et al., 2017) and used in the subsequent analyses (detailed in **Table 1**).

Phylogenetics

A phylogeny was determined from conserved single copy genes present in *P. cactorum* genomes and in *P. idaei* outgroup isolates. Partial and complete single hits from BUSCO searches, using the Alveolata-Stramenopiles obd9 database, were extracted from the 20 sequenced genomes, as well as the publicly available *P. cactorum* 10300 (Armitage et al., 2018) and LV007 genomes (Grenville-Briggs et al., 2017). This led to retention of nucleotide sequences for 230/234 loci, which were aligned using MAFFT

TABLE 1 | Summary of *Phytophthora cactorum* and *Phytophthora idaei* isolates used in this study.

Isolate ID	Genome ID	GenBank accession	Species	Material isolated from	Year	Location	Previously published
P414	Pcac1	NHQK00000000	*P. cactorum*	Strawberry crown	2011	Somerset, UK	
P404	PC116	RCMJ00000000	*P. cactorum*	Strawberry crown	1998	UK	
P415	PC118	RCML00000000	*P. cactorum*	Strawberry crown	2013	UK	
P416	PC119	RCMM00000000	*P. cactorum*	Strawberry crown		UK	
P421	PC129	RCMV00000000	*P. cactorum*	Strawberry crown	2017	UK	
PC13/15	PC122	RCMP00000000	*P. cactorum*	Strawberry crown	2015	UK	
10300	PC110	GCA_003287315.1	*P. cactorum*	Strawberry crown	2006	Norway	Armitage et al., 2018
4032	PC115	RCMI00000000	*P. cactorum*	Strawberry crown		Netherlands	
4040	PC117	RCMK00000000	*P. cactorum*	Strawberry crown		Netherlands	
2003-3	PC114	RCMH00000000	*P. cactorum*	Strawberry crown		Netherlands	
12–420	PC111	RCME00000000	*P. cactorum*	Strawberry crown	2012	Florida, USA	
15–7	PC113	RCMG00000000	*P. cactorum*	Strawberry crown	2015	Florida, USA	
15–13	PC112	RCMF00000000	*P. cactorum*	Strawberry crown	2015	Florida, USA	
11–40	PC127	RCMT00000000	*P. cactorum*	Strawberry fruit	2011	Florida, USA	
17–21	PC128	RCMU00000000	*P. cactorum*	Strawberry fruit	2017	Florida, USA	
62471	PC120	RCMN00000000	*P. cactorum*	Apple	2014	Kent, UK	
P295	PC121	RCMO00000000	*P. cactorum*	Apple (collar rot)	1984	Offham, UK	
R36/14	PC123	RCMQ00000000	*P. cactorum*	Apple	2014	Kent, UK	
LV007[a]		GCA_002081965.1	*P. cactorum*	European Beech	2016	Sweden	Grenville-Briggs et al., 2017
SCRP370	PI125	RCMR00000000	*P. idaei*	Raspberry	1985	Scotland, UK	
SCRP371	PI124	QOKR00000000	*P. idaei*	Raspberry	1986	England, UK	
SCRP376	PI126	RCMS00000000	*P. idaei*	Raspberry	1993	England, UK	

[a]*Only sequence data was used in this study.*

v6.864b (Katoh and Standley, 2013), before being trimmed with trimAl v.1.4.1 (Capella-Gutiérrez et al., 2009). A maximum likelihood tree was determined for each locus using RAxML v.8.1.17 (Liu et al., 2011), with the most parsimonious tree for each locus used to determine an overall consensus phylogeny across all 230 loci using ASTRAL v.5.6.1 (Zhang et al., 2018). The resulting tree was visualised using the R package GGtree v.1.12.4 (Yu et al., 2016).

SNP and Variant Calling

Single Nucleotide Polymorphisms (SNPs), indels and structural variants were identified in reference to the *P. cactorum* P414 genome. Trimmed Illumina reads from each isolate were aligned to the P414 genome using Bowtie2 v2.2.6 (Langmead and Salzberg, 2012), with SNP variants identified using GATK (McKenna et al., 2010; DePristo et al., 2011; Auwera et al., 2013) and indels/structural variants identified using SvABA (Wala et al., 2018). SNPs called by GATK were filtered using VCFtools (Danecek et al., 2011), retaining bi-allelic SNPs with an QUAL $>$ 30, MQ $>$ 40, DP $>$ 10, GQ $>$ 30. SNP calls were also filtered if isolate P414 Illumina reads showed a homozygous polymorphism in reference to the P414 assembly as these represent errors in the assembly rather than SNP variants. Effects of predicted variants on *P. cactorum* gene models were established using SnpEff (Cingolani et al., 2012). Population genetic statistics were calculated from SNP variants using VCFtools and the R package PopGenome (Pfeifer et al., 2014). Structure analysis was performed using FastSTRUCTURE v1.0 (Raj et al., 2014). The program was run with k values between 1 and 6 and number of populations determined where k maximised marginal likelihood. DISTRUCT plots were generated from output meanQ files using R-studio v1.1.453. A SNP distance matrix was made showing the number of variants that differ between isolates. SNP variants were extracted from the final.vcf file as a fasta alignment of concatenated variable sites, containing two sequences per isolate (representing the first and second allele called at each site, respectively). A distance matrix was calculated in Geneious Prime v2020.0 and exported into Microsoft Excel.

Phytophthora spp. Zoospore Production

The production of zoospores was followed from Nellist et al. (2019) and is detailed in the **Supplementary Materials and Methods**. The concentration of zoospores was determined using a haemocytometer and adjusted to 1×10^4, 2×10^4 or 5×10^3 zoospores per mL by diluting with dilute compost extract. The adjusted solution was kept on ice until ready to be used to inoculate plants/unripe fruit.

Pathogenicity Tests on Strawberry Crowns

The virulence of the 18 *P. cactorum* and three *P. ideai* isolates were tested on the crowns of 10 clonal replicates of three cultivars of cultivated strawberry (*F.* × *ananassa*). 'Malling Opal,' an extremely crown rot susceptible cultivar, 'Elsanta' a susceptible cultivar and 'Fenella' a cultivar with good resistance to *P. cactorum*, were screened.

The preparation of plant material is detailed in **Supplementary Materials and Methods** and the inoculation procedure for coldstored strawberry plants was performed as described in Nellist et al. (2019). The data for the ten replicates were averaged and a mean crown rot disease score was used for further analysis. Statistical analyses were performed using R (v3.6.0, "Planting of a Tree"; R Core Team, 2019). A one-way ANOVA was performed to analyse the difference between the pathogenicity of isolates cultured from strawberry on the three cultivars of strawberry.

Pathogenicity Tests on Detached Unripe Strawberry Fruit

Unripe 'Elsanta' strawberry fruit were picked while still white/green in the Summer of 2018. The fruit were then surface sterilised by immersion in a 10% bleach solution and then rinsed twice with dH_2O. The fruit were dried off and two fruit were placed into each sterile 90 mm triple-vented petri dish bottom or lid (Thermo Scientific). The petri dish lids and bottoms with fruit were placed on trays sterilised with 70% ethanol. A sterile 4 mm cork borer was used to bore a shallow hole in the fruit. Zoospores were produced as described above and 100 μL of 5×10^3 zoospore suspension was added into the hole of the fruit. Fruit were screened in three separate experiments with a minimum of eight replicates per isolate screened in experiment. The trays were then sealed in a plastic bag and left in the dark at 20°C. The ratios of colonised to non-colonised fruit were recorded after 7 days.

Pathogenicity Tests on Excised Apple Shoots

Dormant first year growth apple shoots were collected from 'Cox' and 'Gala' in the Winter 2018. The processing of apple shoots was followed from Luberti et al. (2021) and is detailed in the **Supplementary Materials and Methods**. Shoots were assessed for maximum lesion length at 4 weeks by removing the bark around the wound using a scalpel. A digital calliper was used to take measurements and the original wound size, 4 mm, was subtracted from each measurement. A one-way ANOVA was performed to analyse the difference between the pathogenicity of isolates cultured from apple and strawberry fruit on the two cultivars of apple.

In vitro Strawberry Root Pathogenicity Transcriptome Analysis

Transcriptome changes during host infection were investigated through an infection time-course on strawberry roots infected with *P. cactorum* isolate P414. The time-course was performed in the susceptible cultivar 'Emily' and moderately resistant cultivar 'Fenella'; parents of a mapping population used in a previous study (Nellist et al., 2019). Micropropagated plants were produced by GenTech Propagation Ltd. for these experiments. Upon arrival at NIAB EMR, plants were transferred to 120 × 120 × 15 mm, four vent, petri dishes (Corning, Gosselin), half filled with ATS (*Arabidopsis thaliana* salts) media, two plants per plate. ATS media was prepared as described by Taylor et al. (2016). The media were poured into the bottom plate and after it had set, half of the agar was excised with a sterile flat spatula.

Plants were then transplanted so the crown sat on the top of the agar (**Supplementary Figure 1A**). The roots were gently smoothed down, ensuring they were touching the agar. The plates were then sealed with Sellotape and aluminium foil cases were made to surround the agar ensuring a dark environment for the root system (**Supplementary Figure 1B**). The plates were then positioned upright in a growth cabinet (Panasonic MLR-325H) at 22°C, on a 16/8 h, day/night light cycle with a photosynthetic photon flux (PPF) of 150 μmol m^{-2} s^{-1} provided by fluorescent lamps (FL40SSENW37).

Just before inoculation, the micropropagated 'Emily' and 'Fenella' plants were transferred to fresh plates and zoospores were produced as described above. Each root system was inoculated with 1 mL of 2×10^4 zoospore suspension, using a pipette and slowly dripping the suspension over the entire root system. The plates were then sealed with Sellotape, partially recovered with the aluminium foil and were kept flat for 2 h to allow the zoospores to encyst. Mock inoculated (0 h post inoculation) were inoculated with 1 mL of dilute compost extract. The plates were then returned to their upright position until harvested. Root samples were collected at 0 (mock), 6, 12, 24, 48, 72, 96, 120, and 144 h post inoculation (hpi). The root systems were swilled in sterile dH_2O to remove any agar, patted dry and were collected in 2 mL Eppendorf microcentrifuge tubes, flash frozen in liquid nitrogen and stored at −80°C.

Total RNA was extracted from the strawberry roots following a modified version of Yu et al. (2012), over 2 days, detailed in **Supplementary Materials and Methods**. At least 1 μg of root RNA with a RIN score above 7 and with 260/280 and 260/230 ratios above 1.8 were sent to Novogene for sequencing. Strawberry root samples were sequenced to a depth of 50 million reads per sample.

Timepoints for sequencing were selected through the detection of β-tubulin transcripts by Reverse Transcriptase-PCR, using SuperScript™ III Reverse Transcriptase kit with an equal amount of RNA used for each sample. The complementary DNA (cDNA) was then analysed by PCR with 200 μM dNTPs, 0.2 μM of each primer (detailed in **Supplementary Table 1**), 2 μL of cDNA template and 2.5 units of Taq DNA polymerase and the buffer supplied in a 20 μL reaction. Reactions were conducted in a Veriti 96-well thermocycler with an initial denaturation step at 95°C for 30 s, followed by 35 cycles of a denaturation step at 95°C for 30 s, an annealing temperature of 60°C for 30 s and an extension step of 72°C for 30 s. This was followed by a final extension step of 72°C for 5 min and held at 10°C. Products were visualised by gel electrophoresis on a 1% w/v agarose gel at 80 V for 90 min, stained with GelRed. Following this, three biological replicates of samples taken at: 0, 12 and 48 hpi for both 'Emily' and 'Fenella' were sequenced (**Supplementary Figure 3**).

Mycelia of P414 were grown in clarified V8-juice broth as described as above with the addition of 500 μg/mL of ampicillin (Fisher) and 10 μg/mL of rifampicin (Fisher) at 20°C under lab light/dark cycle for 10 days in a shaker incubator set to 200 rpm. The mycelia were washed in sterile dH_2O, vacuum filtered, flash frozen in liquid nitrogen and stored at −80°C. Total RNA was extracted from three biological replicates of flash frozen P414 mycelia using the RNeasy Mini Kit (Qiagen) following the manufacturer's protocol. RNA quality and quantity were assessed as described above and the RNA was sent to The Earlham Institute, UK for sequencing. cDNA library insert sizes were 450–625 bp and were sequenced on Illumina HiSeq4000, using 2x 150 bp PE Reagent Kit. Three barcoded biological replicates of each treatment were pooled and sequenced across multiple lanes. *P. cactorum* mycelia were sequenced to a depth of 25 million reads per sample.

Illumina adapters and low-quality bases were trimmed using fastq-mcf. All RNAseq data were aligned to the *Fragaria vesca* genome v1.1 (Shulaev et al., 2011) using STAR v2.5.3a (Dobin et al., 2013), to remove strawberry reads from the dataset. Read alignment for differential gene expression was performed using Salmon v0.9.1 (Patro et al., 2017) with differential gene expression during infection investigated using DESeq2 (Love et al., 2014). The normalised expression value was represented by applying Fragments per Kilobase of exon model per Million mapped reads (FPKM). All mycelial genes with an FPKM value <1 were adjusted to 1. The Log Fold Change (LFC) in gene expression was calculated using adjusted FPKM values to prevent overprediction of LFC from low-expressed/unexpressed genes under one condition, with all genes with an FPKM value <1 were adjusted 1; LFC = $\log_2$('Emily' 12 hpi FPKM)/$\log_2$(Adjusted Mycelium FPKM). Genes were designated as differentially expressed if they had a DeSeq2 P-adj <0.05 and LFC was ≥2 or was ≤−2. The top 100 expressed genes were investigated further, based on LFC in descending order of 'Emily' at 12 hpi vs. Mycelium. Temporal expression was assessed by identifying those differentially expressed genes with a consistent in peak expression (based upon LFC) in both 'Emily' and 'Fenella' at 12 hpi (early-expressed genes) or 48 hpi (late-expressed genes).

Comparative Genomics of Known Virulence Related Factors

A selection of known RxLR *Avr* gene sequences (amino acid sequence after the signal peptide) were BLASTed against the 22 *Phytophthora* genomes in Geneious Prime v2020.0 and the hits (tBLASTx $E > 1 \times 10^{-10}$) were investigated. Expression of interesting hits were analysed using the *in planta* RNAseq data and interesting candidates were further explored in the representative isolates in the unripe fruit assay.

Strawberry Fruit Reverse Transcription Quantitative Polymerase Chain Reaction Screen

Unripe, green/white "Driscoll® Amesti™" strawberry fruit were used for the pathogenicity time course of three *P. cactorum* isolates; P414, R36/14, and 17-21. The fruit were sterilised and prepared as described above. Zoospores were produced as described above and 100 μL of 5×10^3 zoospore suspension was added into the hole of each fruit. Samples were taken at 0, 36, 48, and 60 hpi. A larger cork borer of 10 mm was used to excise an area around the inoculation point of the fruit, the excised samples were flash frozen in liquid nitrogen and stored at −80°C.

Mycelia of P414, R36/14, and 17-21 were grown in clarified V8-juice broth and harvested as described above. Total RNA

was extracted from the fruit as described above for strawberry roots and from the flash frozen mycelia of P414, R36/14, and 17-21 as described above. The RNA was assessed by the NanoDrop and Qubit 2 as described above. RNA samples were normalised to 720 ng. Reverse transcription was performed on three biological replicates of each *in planta* time point and two biological replicates for the mycelia timepoints with the QuantiTech Reverse Transcription Kit (Qiagen). *In planta* timepoints, for further analysis, were confirmed through the positive identification of *P. cactorum* β-tubulin product by Reverse Transcription-Polymerase Chain Reaction (RT-PCR).

Reverse Transcription quantitative PCR (RT-qPCR) was then performed in a CFX96^{TM} Real-Time PCR detection system (BioRad) in 10 μL reactions of: 5 μL of 2x qPCRBIO SyGreen Mix Lo-Rox (PCR Biosystems), 2 μL of a 1:5 dilution of the cDNA sample in dH_2O and 400 nM of each primer (detailed in **Supplementary Table 1**). Due to the presence of primer dimers, an additional step at the end of the reaction was added to measure the fluorescence at a temperature greater than the melting temperature of the primer dimers (Ball et al., 2003). The reaction was run with the following conditions: 95°C for 2 min, 40 cycles of 95°C for 5 s, 62°C for 20 s and 80°C for 5 s. This was followed by 95°C for 10 s, and a 5 second step ranging from 65 to 95°C by 0.5°C every cycle to generate melt curves. At least two technical replicates for each sample were performed and the melt curves were analysed to ensure the correct product was detected. Relative gene expression was calculated using the efficiency corrected method, which determines the relative gene expression ratio based on the real-time PCR efficiencies and the cycle quantification value (C_q; Pfaffl, 2001):

$$ratio = \frac{(E_{\text{target}})^{\Delta CP_{\text{target}}(control-sample)}}{(E_{\text{ref}})^{\Delta CP_{\text{ref}}(control-sample)}}$$

Expression values were calculated as the mean of the three biological replicates and the standard error of the mean was calculated and plotted. A pooled sample of all cDNA was used as an inter-plate control (IPC) on all plates using primers for β-tubulin (Pcac1_g23639; **Supplementary Table 1**). Genes of interest were normalised to two endogenous reference genes (**Supplementary Table 1**), a ribosomal 40S protein (Pcac1_g24902) and a protein of the BAR-domain family, Pc_WS41, Pcac1_g27577 (Yan and Liou, 2006) and were plotted relative to the expression of the gene of interest in mycelia.

RESULTS

Phytophthora cactorum Isolates Show Specialisation to Strawberry Crowns and Apple

A clear difference in pathogenicity on different plant tissues was observed between the isolates from the different *P. cactorum* pathotypes (**Figure 1**, **Supplementary Figure 2**). Variation in pathogenicity on strawberry crowns was observed between the *P. cactorum* isolates cultured from strawberry ($p = 0.02$; **Figure 1A**). All strawberry CR isolates were able to cause disease in strawberry crowns to varying degrees on the three different strawberry cultivars (**Figure 1A**). Of the two strawberry LR isolates, 11–40 was able to cause disease in both 'Malling Opal' and 'Elsanta,' whereas 17-21 was only able to cause disease in the very susceptible 'Malling Opal' (**Figure 1A**). None of the apple isolates were able to cause disease in the strawberry crowns (**Figure 1A**). Variation in pathogenicity on excised apple shoots was observed between the *P. cactorum* isolates cultured from apple and strawberry LR ($p = 0.03$; **Figure 1B**). Apple isolate R36/14 was found to be the most pathogenic on the

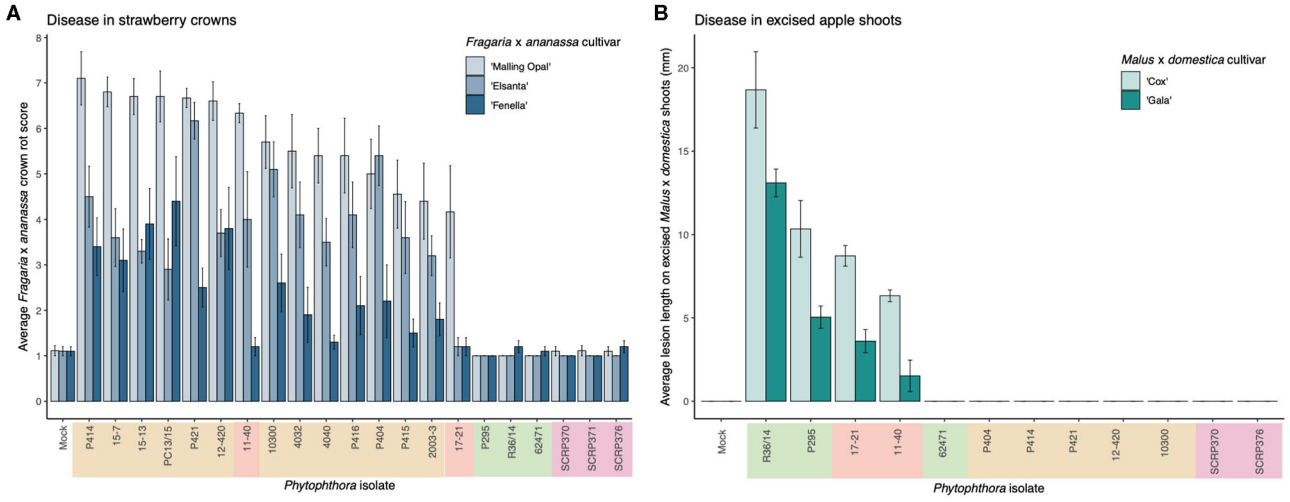


FIGURE 1 | *Phytophthora cactorum* isolates show specialisation to strawberry crown and apple tissue. **(A,B)** Virulence of *P. cactorum* and *P. idaei* isolates on strawberry crowns and excised apple shoots by artificial inoculation of zoospores and mycelium, respectively. The colour behind the isolate names denotes what plant tissue they were cultured from; beige, strawberry crowns, red, strawberry fruit, green, apple bark and purple, raspberry (*P. idaei*). **(A)** Virulence of *Phytophthora* isolates on three *Fragaria* x *ananassa* cultivars, 'Malling Opal,' 'Elsanta,' and 'Fenella.' Data are the mean of ten biological replicates ± se. **(B)** Virulence of *Phytophthora* isolates on two *Malus* × *domestica* cultivars, 'Cox' and 'Gala.' Data are the mean of six biological replicates ± se.

TABLE 2 | Assembly statistics for 17 *Phytophthora cactorum* and three *Phytophthora idaei* isolates sequenced in this study.

Features		*Phytophthora cactorum*													*Phytophthora idaei*		
		Fragaria × *ananassa* crowns									*Fragaria* × *ananassa* fruit		*Malus* × *domestica*			*Rubus idaeus*		
		P414	12–420	15–13	15–7	2003–3	4032	4040	P404	P415	P421	PC13/15	11–40	17–21	62471	P295	R36/14	SCRP370	SCRP371	SCRP376
Coverage		76.49	25.60	55.63	59.71	69.29	53.14	67.21	69.52	86.64	106.94	77.83	68.90	78.86	72.21	39.29	56.09	40.91	71.89	49.19
No. of contigs		194	5,509	5,421	5,457	5,467	6,726	5,326	20,136	5,695	5,485	5,278	4,571	4,452	6,173	5,088	5,952	5,356	4,720	5,313
Assembly size (Mb)		66.0	59.9	60.7	60.8	60.8	61.6	60.7	75.5	60.8	60.5	60.0	61.1	59.7	59.8	61.3	60.7	60.5	61.7	60.6
Longest contig (bp)		1,749,730	272,591	358,283	300,516	353,051	299,428	238,735	346,862	313,236	251,913	346,832	288,536	347,645	285,881	357,485	345,260	247,223	295,942	230,545
N50 (bp)		645,375	37,498	41,248	43,376	40,273	38,202	39,295	36,598	40,742	34,103	39,779	43,828	49,722	36,408	48,651	36,465	39,576	46,381	39,406
Masked (% bp)		29.2	24.9	26.2	26.0	26.2	26.1	26.1	18.4	26.5	26.0	25.9	24.5	25.1	25.6	25.6	26.2	27.0	26.6	27.3
BUSCO	C	224	228	227	228	228	228	228	227	228	226	228	227	227	226	227	227	230	230	229
	% present	95.7	97.4	97.0	97.4	97.4	97.4	97.4	97.0	97.4	96.6	97.4	97.0	97.0	96.6	97.0	97.0	98.3	98.3	97.9

Isolate P414 represents results from long read sequencing technologies whereas other assemblies are a result of short read sequencing technology. Information on sequencing coverage, assembly size, percentage of bases repeatmasked are shown. Assembly completeness statistics, assessed using Alveolata-Stramenopiles Benchmarking Universal Single-copy Orthologue (BUSCO) genes, detail the number of complete (C) and total percentage of 234 core Eukaryotic genes present in assemblies.

shoots (**Figure 1B**). *P. cactorum* isolate 62471 was shown to be pathogenic on apple seedlings in a previous screen in 2018 (data not shown). Of the two strawberry LR isolates, 17–21 was more pathogenic on the apple shoots than 11–40 (**Figure 1B**). None of the strawberry CR isolates tested were able to cause disease in apple shoots (**Figure 1B**). The two strawberry LR isolates appear to have a broader host range than either strawberry CR or apple isolates and are able to cause disease in both strawberry crowns and apple shoots (**Figure 1**). All representative isolates from the three pathotypes of *P. cactorum* were able to colonise strawberry fruit (**Supplementary Figure 2**). Isolate 62471 appeared to be the weakest apple isolate as it caused the lowest percentage infection in the strawberry fruit over the three experiments (**Supplementary Figure 2**), coinciding with its weak pathogenicity on apple tissue. No disease symptoms were recorded in the strawberry tissue or apple tissue when challenged with the *P. idaei* isolates (**Figure 1**, **Supplementary Figure 2**). Although, it should be noted that the *P. idaei* isolates were also tested on raspberry fruit but the results were inconclusive.

Generation of an Improved Contiguous *P. cactorum* Genome

Here we present the best assembly to date of the plant pathogen *P. cactorum*. The SMRT data for the strawberry CR isolate P414 yielded an assembly of 66 Mb in 194 contigs. The other *P. cactorum* isolates, with the exception of P404, yielded *de novo* Illumina assemblies of 59.7–61.6 Mb in 4,452–6,726 contigs (**Table 2**), with isolate P404 larger and in a greater number of contigs, totalling 75.5 Mb in 20,136 contigs. The *P. idaei* genomes were a similar size to *P. cactorum* assemblies, 60.4–60.6 Mb in 4,720–5,356 contigs. Gene space between all assemblies was comparable, with 224–230 of 234 (95.7–98.3%) BUSCO genes both present and complete in the assemblies (**Table 2**), which were comparable to previous *Phytophthora* spp. sequencing projects, 91.5–94.4% for *P. cinnamomi* (Longmuir et al., 2018).

Whole Genome Phylogeny Supports Resolution Between *P. cactorum* Pathotypes

A consensus phylogeny of 230 conserved single copy genes from the 22 *Phytophthora* isolates showed clear resolution between species, with *P. cactorum* and *P. idaei* isolates resolved into distinct clades (**Figure 2**). Resolution was also shown within *P. cactorum*, with the 13 strawberry CR isolates, including the previously sequenced CR isolate 10300, present in a distinct clade from the three apple isolates. The two strawberry LR isolates were placed into different clades. Strawberry LR isolate 11–40, which was more virulent on strawberry crowns, was placed in the same clade as the strawberry CR isolates. Whereas, strawberry LR isolate 17–21, which was more virulent on apple, was placed in the same clade as the apple isolates. Interestingly, the publicly available *P. cactorum* isolate from *F. sylvatica* was genetically distinct from all other *P. cactorum* isolates (**Figure 2**). Within the strawberry clade, no evidence was observed for isolates being associated with geographical distribution, with isolates

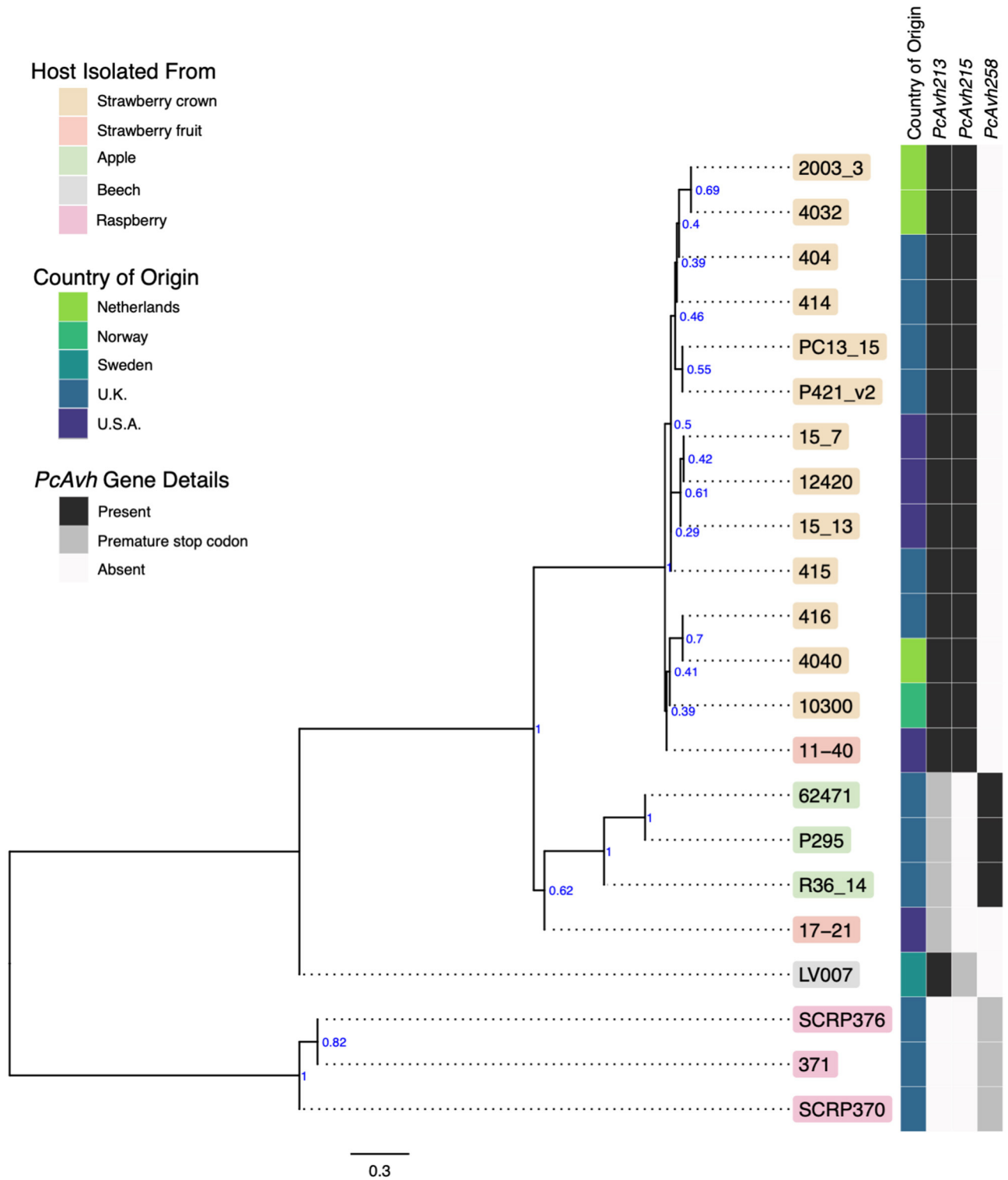
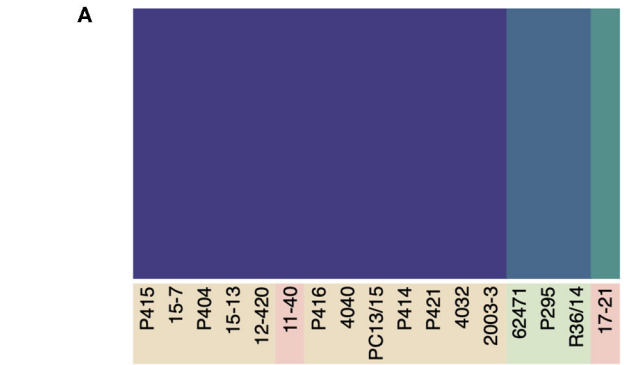
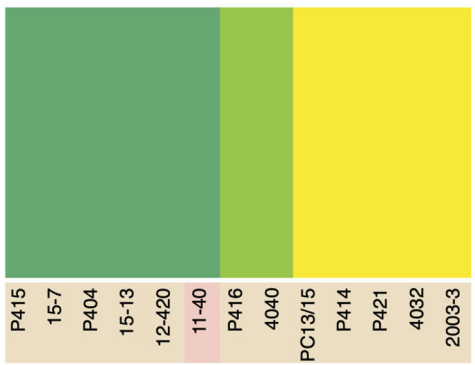


FIGURE 2 | Apple and strawberry crown rot *Phytophthora cactorum* isolates form distinct clades. Maximum parsimony consensus phylogeny of 230 Alveolata-Stramenopiles Benchmarking Universal Single-Copy Orthologue (BUSCO) genes from 22 *Phytophthora* isolates. Branch support shown in blue. Details of the host species isolated from, country of origin and presence/absence of a *Phytophthora sojae Avh32* homologue (*PcAvh215*) and *Phytophthora infestans Avr3a* homologue (*PcAvh258*) and *PcAvh213* are shown in the heatmap.





C

		P415		15-7		P404		15-3		12-420		11-40		P416		4040		PC13/15		P414		P421		2003-3	
		a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b
P415	a		65	22	103	24	92	24	89	23	93	31	96	983	1057	984	1058	564	629	553	617	552	615	552	616
	b			87	38	89	27	89	24	88	28	96	31	1046	992	1047	993	629	568	618	556	617	554	617	555
15-7	a				81	22	90	20	85	9	79	27	92	979	1053	980	1054	560	625	549	613	548	611	548	612
	b					103	41	101	36	90	30	108	43	1058	1004	1059	1005	641	580	630	568	629	566	629	567
P404	a						68	24	89	23	93	31	96	983	1057	984	1058	564	629	553	617	552	615	552	616
	b							92	27	91	31	99	34	1049	995	1050	996	632	571	621	559	620	557	620	558
15-3	a								65	21	91	15	80	981	1055	982	1056	562	627	551	615	550	613	550	614
	b									86	26	80	15	1044	990	1045	991	627	566	616	554	615	552	615	553
12-420	a										70	28	93	980	1054	981	1055	561	626	550	614	549	612	549	613
	b											98	33	1048	994	1049	995	631	570	620	558	619	556	619	557
11-40	a												65	988	1062	989	1063	569	634	558	622	557	620	557	621
	b													1051	997	1052	998	634	573	623	561	622	559	622	560
P416	a														76	3	79	1055	1118	1044	1106	1043	1104	1043	1105
	b															79	3	1131	1070	1120	1058	1119	1056	1119	1057
4040	a																76	1056	1119	1045	1107	1044	1105	1044	1106
	b																	1132	1071	1121	1059	1120	1057	1120	1058
PC13/15	a																		67	21	87	20	85	26	92
	b																			88	26	87	24	93	31
P414	a																				66	13	78	15	81
	b																					79	16	81	19
P421	a																						65	14	80
	b																							79	17
2003-3	a																								66
	b																								

FIGURE 3 | *Phytophthora cactorum* isolates split into three populations with pre-dominant asexuality. Population analysis of high quality, biallelic SNP sites split the *P. cactorum* isolates into three populations, admixture of variants was not observed within *P. cactorum*, or within a single lineage of *P. cactorum* indicating predominant asexuality. DISTRUCT plot of fastSTRUCTURE (Raj et al., 2014) results carried out on **(A)** all sequenced isolates of *P. cactorum* and **(B)** strawberry lineage isolates. Data was best described by three "populations" in both datasets. **(C)** SNP calling was performed in relation to reference isolate P414 for the *P. cactorum* strawberry lineage. The number of differing sites are shown between the two haplotypes from each isolate; a, representing only homozygous variants and b, representing both homozygous and heterozygous variants in relation to the P414 reference genome.

from the U.K., Netherlands, Norway and U.S.A., observed to be distributed throughout the clade.

Population Structure Reflects Predominant Asexuality Within Diverging *P. cactorum* Lineages

Population structure within *P. cactorum* was investigated using SNP variants predicted in relation to the strawberry CR isolate P414. *P. cactorum* SNP data was found to be best described by three non-recombining populations representing the strawberry lineage, apple lineage and the LR isolate 17–21, respectively (**Figure 3A**). Proportions of shared variants between these populations was <0.1% in each isolate, indicating a lack of recombination between populations. Subpopulation structure also showed indications of asexuality when restricted to isolates within the strawberry lineage (**Figure 3B**). Isolates within the three subpopulations of the strawberry lineage showed high genetic identity to one another, with the number of heterozygous sites observed within a diploid individual being close to the genetic distance between two individuals from the same population (**Figure 3C**). For example, in the subpopulation consisting of 4040 and P416, 44 sites were found to be heterozygous within isolate 4040, whereas the total number of SNPs that showed variability between isolate 4040 and P416 (including heterozygous and homozygous sites) was observed to be 79 (**Figure 3C**). This was in contrast to the number of SNPs differing between isolates between subpopulations, where isolate 4040 differed by a total of 1,054 SNPs to its next closest isolate, 15–7, from another subpopulation (**Figure 3C**).

P. cactorum Possesses an Expanded Effector Repertoire in Comparison to *P. idaei*

The predicted proteome of *P. cactorum* CR isolate P414 totalled 29,913 proteins encoded by 29,552 genes. Additional isolates

were predicted to carry a similar number of proteins 25,449–29,955 and genes 24,856–29,124, with the exception of P404 (**Table 3**). Within *P. cactorum*, strawberry and apple isolates had similar numbers of predicted genes and effectors. Despite the larger number of predicted genes, P404 had a comparable number of predicted RxLR, CRN, and apoplastic effectors to the other *P. cactorum* isolates. However, *P. idaei* isolates carried a reduced predicted effector repertoire (**Table 3**), with fewer secreted carbohydrate active enzymes (CAZYmes), CRNs, Elicitins, necrosis-inducing proteins (NLPs), glucanase inhibitors and kazal protease inhibitors predicted than *P. cactorum* isolates.

Orthology Analysis Identifies Gene Expansion and Contraction Associated With Phylogenetic Lineage

The total set of 157,038 predicted proteins from the 20 sequenced isolates, as well as the proteome of *P. cactorum* isolate 10300 from Armitage et al. (2018), were clustered into 22,572 orthogroups. Orthogroups showing a consistent pattern of expansion/contraction by phylogenetic clade were identified (**Figure 4**). This allowed investigation into expansion/contraction events associated with the strawberry CR and apple lineages of *P. cactorum*.

The 45 orthogroups expanded in the strawberry CR lineage (Branch A; **Figure 4**), represented a total of 65 genes from P414 (**Supplementary Table 2**). This included a number of potential effector candidates, notably two RxLRs (*Pcac1_g24384* and *Pcac1_g22827*) and an additional secreted protein (*Pcac1_g6287*). BLAST searches confirmed the absence of *Pcac1_g24384* in the apple lineage to be the result of these regions being absent from assemblies rather than genes not being predicted in those genomes. PCR of gDNA from R36/14 and 17-21 did not detect the gene. BLAST searches revealed that isolates in the apple lineage have a region that is homologous to *Pcac1_g22827*, however, there is an indel present in all four isolates. There is an additional G at base 99, in comparison to the strawberry CR lineage, resulting in a frame shift mutation and a premature stop codon at amino acid (aa) 38 and the gene not being predicted in these genomes. The strawberry CR lineage was found to have contracted across 21 orthogroups (Branch A; **Figure 4**), representing 33 genes in P414 (**Supplementary Table 2**). This included three RxLRs (*PC123_g16852*, *PC123_g26877* and *PC123_g27632*) and two additional secreted proteins (*PC123_g10425* and *PC123_g25979*) with effector-like structure that were lost in relation to the wider phylogeny.

The apple isolates and 17–21 lineage harboured greater diversity, with 119 orthogroups expanded (Branches B_0, B_1; **Figure 4**), representing a total of 241 genes (**Supplementary Table 2**). This included five RxLR (*PC123_g15654*, *PC123_g17462*, *PC123_g19522*, *PC123_g24792*, and *PC123_g25079*) and two CRN (*PC123_g21108* and *PC123_g24736*) candidates, as well as seven secreted proteins (*PC123_g10425*, *PC123_g10510*, *PC123_g11377*, *PC123_g14333*, *PC123_g24080*, *PC123_g27245*, and *PC123_g28487*), six of which had an effector-like structure (**Supplementary Table 2**). BLAST searches confirmed the absence of these genes in the strawberry CR lineage to be a result of these regions being absent from assemblies rather than genes not being predicted in those genomes. A single RxLR (*PC123_g19522*) and CRN (*PC123_g24736*) candidate were identified unique to apple isolates (Branch B_1; **Figure 4**). PCR of gDNA of *PC123_g19522* (hereafter *PcAvh258*; all RxLR homologues summarised in **Supplementary Table 3**) confirmed the absence of this gene in the strawberry isolates P414 and 17–21. *PcAvh258* was found to have 58% pairwise aa homology (downstream of the signal peptide, aa 24–139) to *P. infestans Avr3a* (GenBank AEH27535.1; aa 22-147) (Armstrong et al., 2005). Further investigation of *PC123_g24792* noted a difference between the apple isolate homologues and that in 17–21. The homologue in the three apple isolates was truncated (69 aa; hereafter *PcAch246t*) compared to *PC128_g25726* in 17-21 (195 aa; hereafter *PcAvh246*). A non-synonymous SNP introduced at G210A resulted in a stop codon.

Contraction of 28 orthogroups was observed in the apple lineage (Branches B_0, B_1; **Figure 4**), representing 45 genes. This included only one RxLR candidate (*Pcac1_g13631*) and three additional secreted proteins (*Pcac1_g3068*, *Pcac1_g3069* and *Pcac1_g25117*), indicating a substantial increase in the effector complement of this lineage.

Overall, these results show that the apple lineage within *P. cactorum*, harbours greater diversity in effector complement than the strawberry CR lineage. RxLRs and CRNs were represented in the expanded and contracted gene families, as well as other unannotated proteins with an effector-like structure. However, other commonly observed *Phytophthora* pathogenicity factors such as secreted CAZYmes, elicitins and protease inhibitors were notably absent from these groups.

Polarising of SNP and Indel Variants Identifies Putative Host Specialisation Events in Effectors

Further variants were determined through identification of SNP, indel and small structural variants (insertions and duplications) from all sequenced isolates in comparison to strawberry CR isolate P414 (**Table 4**). Polarising of non-synonymous SNPs and indels to the outgroup *P. idaei* allowed identification of those variants that differed at the species level (private to *P. cactorum*), at the pathotype level (private to apple or strawberry CR isolates), or at the population level. Those variants at the pathotype level were investigated to identify potential signatures of host adaptation. In total, variants were observed in 21 RxLR and 12 CRN genes at the pathotype level (**Supplementary Table 4**). Of the RxLRs, eight genes contained non-synonymous variants unique to strawberry isolates and 13 unique to apple isolates. Of the CRNs, six genes contained non-synonymous variants unique to strawberry isolates and eight unique to apple isolates (with two genes containing unique variants in both), which in addition to the gene family expansion/contractions described above potentially represent host adaptation events or determinants of host boundaries between pathotypes, or simply functionally neutral mutations fixed due to drift.

TABLE 3 | Predicted gene models and effectors.

Features		*Phytophthora cactorum*																*Phytophthora idaei*		
		Fragaria × *ananassa* crowns											*Fragaria* × *ananassa* fruit		*Malus* × *domestica*			*Rubus idaeus*		
		P414	**12–420**	**15–13**	**15–7**	**2003–3**	**4032**	**4040**	**P404**	**P415**	**P416**	**P421**	**PC13/15**	**11–40**	**17–21**	**62471**	**P295**	**R36/14**	**SCRP370**	**SCRPP371**	**SCRP376**
Genes		29,552	25,444	25,942	25,855	28,950	26,062	28,670	34,978	26,042	28,623	25,566	25,771	28,046	27,869	28,767	25,627	29,124	27,354	24,856	24,924
Proteins		29,913	26,265	26,779	26,646	29,811	26,859	29,557	35,812	26,850	29,425	26,155	26,641	28,851	28,599	29,620	26,368	29,955	27,951	25,449	25,557
Secreted		1,887	1,658	1,686	1,668	1,758	1,707	1,758	2,056	1,719	1,770	1,345	1,697	1,723	1,790	1,771	1,666	1,799	1,610	1,540	1,522
EffectorP		507	472	470	475	525	502	522	660	499	537	381	484	511	532	520	476	543	481	456	443
CWDE		281	222	234	220	222	222	225	261	224	219	215	227	222	237	233	229	217	200	194	195
RxLR		158	134	134	136	136	139	137	135	151	144	135	143	132	144	146	142	146	146	144	138
CRN		127	85	89	88	74	94	88	84	93	88	74	91	70	73	74	93	97	69	59	62
TFs		1,215	904	913	921	1,011	924	978	1,064	946	1,007	861	932	968	952	1,003	928	1,038	905	884	862
MAMPs	Elicitin	43	43	42	46	44	43	47	43	46	43	43	44	44	45	44	44	43	37	32	38
	TGA	14	13	13	13	13	13	13	13	12	13	11	13	12	13	13	13	12	12	12	13
Cutinase	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	2	3	2	2
NLP	22	29	29	28	28	28	26	23	29	27	27	26	24	32	32	17	32	11	19	14	14
Pcf	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	4	4
GI	21	15	16	18	14	16	15	15	17	16	14	15	18	17	15	19	13	8	10	11	11
Protease inhibitors	Kasal	17	17	17	17	16	16	15	16	17	16	15	16	17	16	16	16	16	13	15	14
	Cath.	3	3	3	3	3	3	4	4	3	3	3	3	3	3	3	4	3	3	3	3
	Cyst.	2	1	1	1	2	1	2	2	1	2	1	1	2	1	2	1	2	1	1	1

Total number of gene models and the predicted proteins that they encode are shown. Numbers of secreted proteins, secreted cell wall degrading enzymes (CWDE), RxLR and crinkler (CRN) family cytoplasmic effectors are shown, along with predicted transcription factors (TFs). Host defence triggering (MAMP) family sterol binding (Elicitn) and transglutanimase (TGA) proteins are shown as well as apoplastic effector families including necrosis inducing proteins (NLP), Phytophthora cactorum factor (PcF), and protease inhibitors (kazal-, cathepsin-, and cystatin-acting).

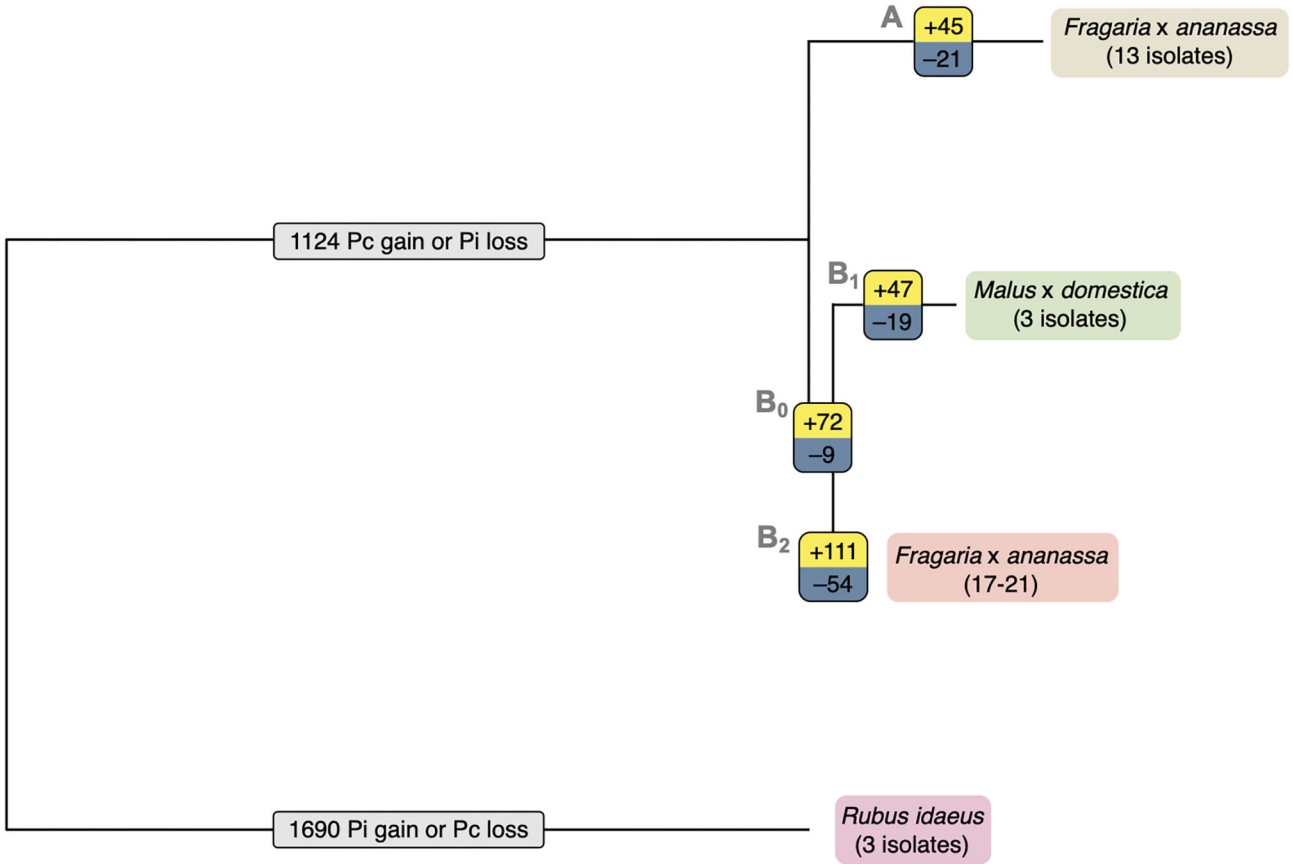


FIGURE 4 | Expansion and contraction of orthogroups indicates possible roles in *P. cactorum* host specialisation. Expansion and contraction of orthogroups plotted onto simple phylogeny. The number of expanded orthogroups of each branch are shown in yellow, the number of contracted orthogroups of each branch are shown in blue and the numbers of orthogroups with unsure direction of expansion or contraction are shown in light grey. The strawberry crown rot lineage (branch A) includes isolates P414, P404, P415, P416, P421, 4032, 4040, 12–420, 2003–3, 10300, 15–7, 15–13, and PC13/15. The apple lineage (branch B_1) includes isolates R36/14, P295, and 62471.

Putative Effectors Are Highly Represented in DEGs During Infection of Strawberry

RNAseq analysis of *P. cactorum* (P414) infecting strawberry at 12 and 48 hpi showed predicted gene models of putative effectors were upregulated during strawberry infection. Differential gene expression was calculated between mycelium, 12 and 48 hpi in both 'Emily' and 'Fenella' cultivars and between 12 and 48 hpi timepoints for both cultivars. This allowed identification of early and late expressed transcripts and of the remaining transcripts, identification of those up- and down-regulated *in planta*. In total 9,178 transcripts with LFC >2 were identified. This equated to 34% of the total transcripts predicted in the genome (**Table 5**). Putative apoplastic and cytoplasmic effectors were overrepresented within the DEGs with 43–76% of candidates differentially expressed in the experiment. Of these, many secreted CAZYme and RxLR candidates showed temporal expression, showing differential expression only the early or late timepoint (160 and 63 transcripts, respectively), with CAZYmes showing a greater number of late expressed candidates and RxLRs showing greater numbers of early expressed candidates (**Table 5**, **Figure 5A**). Furthermore, transglutaminase candidates and Kazal-type protease inhibitors were expressed during early infection, whereas NLP candidates showed a bias towards later infection. Two homologues of *P. infestans* INF1 showed consistent expression across the timepoints (**Supplementary Table 4**; *Pcac1_g22873* and *Pcac1_g22879*). Interestingly, effectors from each category were identified as down-regulated *in planta*, particularly cytoplasmic CRN effectors, of which 69 were down-regulated at both 12 and 48 hpi (**Table 5**, **Figure 5B**). In total, of the 158 putative RxLR effectors identified in P414, just over half, 86 were not expressed or showed low-expression *in planta* (with FPKM values <20) in the RNAseq experiment.

The Most Upregulated Genes *in planta* Include a Broad Range of Effector Candidates

Genes involved in initial establishment of infection were further investigated through ranking transcripts by LFC at 12 hpi in comparison to mycelium in the susceptible host 'Emily' (all

TABLE 4 | Variant calls vs. the reference P414 genome.

	P. cactorum					*P. idaei*
	P414	**CR isolates**	**Apple isolates**	**LR 11–40**	**LR 17–21**	**Raspberry isolates**
Total SNPs	67	1,536	26,333	631	20,482	306,259
Gene SNPs	46	844	13,787	357	10,970	165,226
CDS SNPs	30	748	12,318	304	9,701	147,237
Non-syn/syn SNPs						
Total	17/13	467/281	6,400/5,918	182/122	4,481/5,220	75,340/71,897
Busco CEGs	0/0	4/11	77/119	2/7	68/90	978/1,453
RxLR	0/0	0/0	15/9	0/0	14/3	299/137
CRN	0/0	0/0	6/2	0/0	7/3	74/35
Total InDels	712	1,760	11,383	1,057	9,200	63,362
Gene InDels	162	394	2,249	241	1,791	14,005
CDS InDels	131	316	1,659	191	1,299	9,282
Busco CEG InDels	0	1	6	0	6	64
RxLR InDels	0	2	12	1	14	81
CRN InDels	2	5	12	2	6	32
Total SVs	0	12	16	0	16	92
Gene SVs	0	8	9	0	6	33
CDS SVs	0	8	9	0	6	29
Busco CEG SVs	0	0	0	0	0	0
RxLR SVs	0	0	0	0	0	0
CRN SVs	0	0	0	0	0	0

Numbers of SNPs, insertion/deletion events (InDels) and structural variants predicted in relation to strawberry CR P414. Variants leading to changes within gene models and leading to coding changes in CDS are shown. SNP calls within CDS include the numbers of synonymous/non-synonymous variants.

TABLE 5 | Expression profile of effector candidates.

	Early induced	**Late induced**	**Upregulated *in planta***	**Downregulated *in planta***	**Other DE[a]**	**Total in genome**	**% DEGs of total[b]**
Total transcripts	437	809	2,613	5,319	1,117	29,913	34.4
Secreted	80	141	442	351	91	1,887	58.6
MAMP: elicitin	4	6	23	13	5	67	76.1
MAMP: transglutaminase	4	0	4	3	1	18	66.7
Apoplastic: secreted CAZYmes	19	36	105	40	14	281	76.2
Apoplastic: cutinase	1	0	3	0	0	6	66.7
Apoplastic: glucanase inhibitor	1	1	10	4	0	30	53.3
Apoplastic: NLP	2	6	10	3	0	35	60.0
Apoplastic: phytotoxin	0	0	1	1	0	3	66.7
Apoplastic: protease inhibitor (cathepsin)	0	0	0	2	1	4	75.0
Apoplastic: protease inhibitor (cystatin-like)	0	0	2	0	1	5	60.0
Apoplastic: protease inhibitor (Kazal-type)	5	1	6	4	0	23	69.6
Cytoplasmic: RxLR	9	6	48	4	1	158	43.0
Cytoplasmic: CRN	1	1	2	69	2	127	59.1

Numbers of differentially expressed (DE) transcripts, summarised by effector family, showing DE at 24 hpi (early induced) or 48 hpi (late induced) in both Fragaria x ananassa 'Emily' and 'Fenella' cultivars, or upregulated at both timepoints or downregulated at both timepoints in planta.

[a] *DE, Differential expression.*

[b] *DEGs, Differentially expressed genes.*

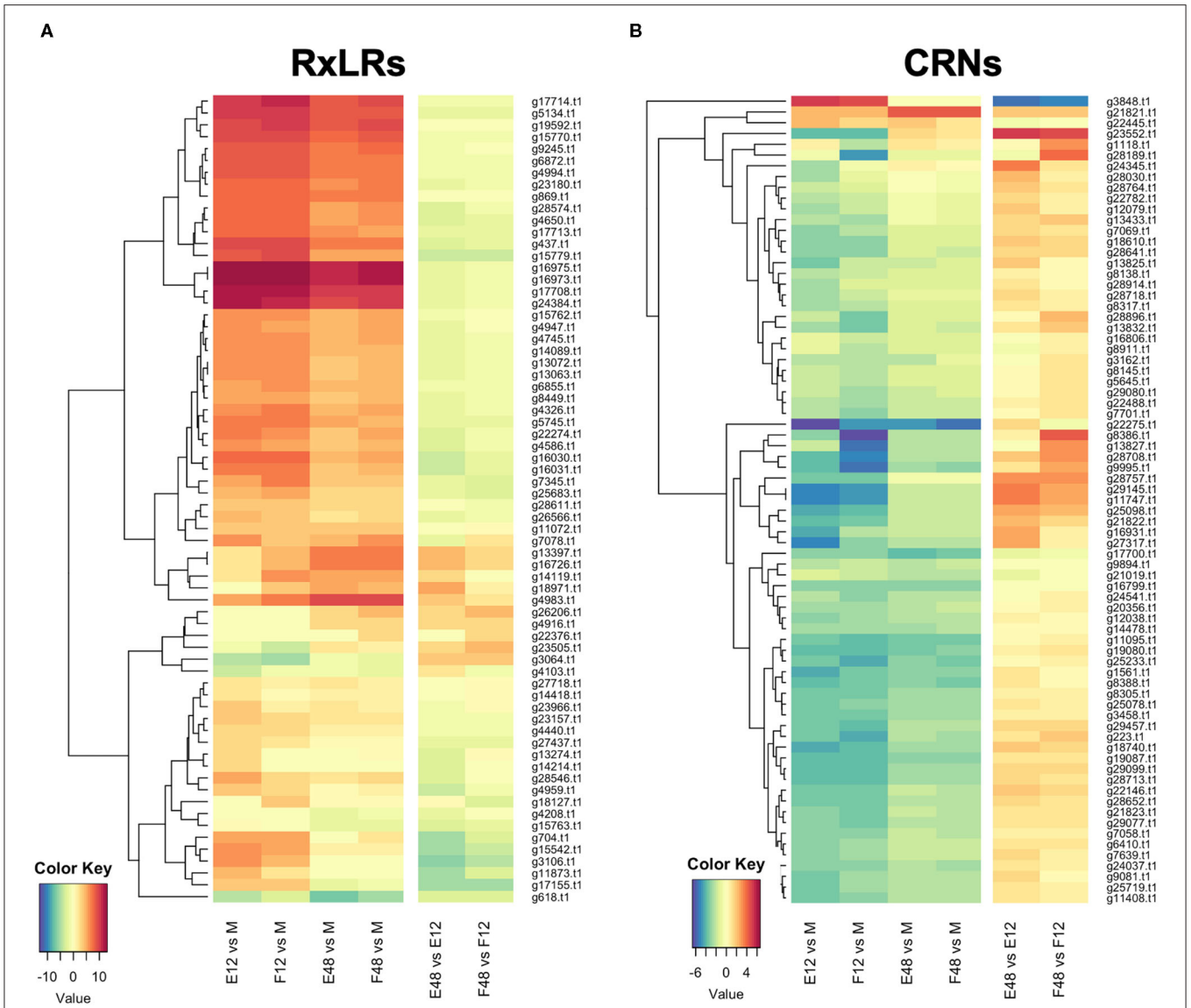


FIGURE 5 | Expression pattern of effectors observed upon infection of both susceptible ('Emily') and resistant ('Fenella') strawberry roots. RxLR and crinkler (CRN) genes clustered by expression profile. All gene IDs related to the P414 genome (*Pcac1_*). E, 'Emily;' F, 'Fenella;' M, mycelia; 12, 12 hours post-inoculation (hpi); 48, 48 hpi. **(A)** RxLR genes shown in dark red were in the top 25 expressed genes at 12 hpi ranked by log fold change (LFC), **(B)** the time points investigated don't appear to have captured differential expression of CRNs.

data can be found in **Supplementary Table 4**). In the top 100 ranked genes (LFC values of 16.8–8), secreted proteins were highly represented with 56 present, including 15 RxLR candidates (along with two additional candidates, that carried RxLR motifs but lacked EER motifs), 13 secreted CAZymes (AA7, CBM63, CE8, GH12, GH28, and PL3 families), four NLP candidates, three kazal-type protease inhibitors (one of which was homologous to *P. infestans* EPI10), two glucanase inhibitors and *P. cactorum* factor (PcF) homologue phytotoxin.

Focusing on the top 25 ranked genes (summarised in **Table 6**), six RxLR candidates were upregulated upon infection of strawberry (*Pcac1_g16973*, *Pcac1_g16975*, *Pcac1_g17708*, *Pcac1_g24384*, *Pcac1_g5134*, and *Pcac1_g17714*) and an additional one (*Pcac1_g998*), which carried an RxLR motif without an EER motif. *Pcac1_g24384* (hereafter *PcAvh215*) was of particular interest as it was identified as unique to *P. cactorum* strawberry CR isolates and highly expressed *in planta*. Subsequent RT-qPCR analysis in strawberry fruit supported the findings of the RNAseq timepoints and showed expression of *PcAvh215* (**Figure 6A**). Notably, this was the only example of a gene private to strawberry CR isolates in the top 100 ranked transcripts. Investigation into this gene showed it to be a homologue of *P. parasitica* XM_008912329 and *P. sojae Avh32* (JN253712; paralogue to *Avh6;* Wang et al., 2011) with 86 and

TABLE 6 | The most upregulated genes in 'Emily' at 12 h post infection include a broad range of effector candidates.

LFC	Expression pattern	Transcript ID	Contig	Orthogroup	Presence by group	Presence variation at level:	Non-Syn SNP/indel variation at level	Secretion evidence	Function annotation
16.8	Early	g20321.t1	Contig_57	Orthogroup89	CR(14)LR(1)Md(3)Ri(3)				
15.6	Early	g2828.t1	Contig_4	Orthogroup15102	CR(10)LR(1)Md(2)Ri(0)	*		Phobius	Coil domain
15.5	Early	g9289.t1	Contig_18	Orthogroup16991	CR(7)LR(0)Md(0)Ri(0)	*			
15.3	Early	g12985.t1	Contig_29			*			Coil domain
14	Early	g7826.t1	Contig_14	Orthogroup15950	CR(9)LR(0)Md(1)Ri(0)	*	Species		
13.6		g12191.t1	Contig_26	Orthogroup359	CR(14)LR(1)Md(3)Ri(3)			SignalP; Phobius	EffectorP; Jacalin-like lectin
13.1		g998.t1	Contig_2	Orthogroup774	CR(14)LR(1)Md(3)Ri(3)			SignalP; Phobius; TM domain	RxLR (no EER motif); EffectorP
13.1	Early	g24044.t1	Contig_79	Orthogroup14845	CR(11)LR(1)Md(2)Ri(0)	*	Species		
13		g16973.t1	Contig_43	Orthogroup12686	CR(14)LR(1)Md(3)Ri(0)	*P. cactorum* private		SignalP; Phobius	RxLR; EffectorP
13		g16975.t1	Contig_43	Orthogroup12686	CR(14)LR(1)Md(3)Ri(0)	*P. cactorum* private		SignalP; Phobius	RxLR; EffectorP
12.2		g12202.t1	Contig_26	Orthogroup359	CR(14)LR(1)Md(3)Ri(3)		Species	SignalP; Phobius	EffectorP; Jacalin-like lectin
11.9		g17708.t1	Contig_46	Orthogroup9680	CR(14)LR(1)Md(3)Ri(3)		Species	SignalP; Phobius	RxLR
11.5		g26017.t1	Contig_95	Orthogroup11598	CR(14)LR(1)Md(3)Ri(3)		Species	SignalP; Phobius	
11.4		g24384.t1	Contig_81	Orthogroup14979	CR(14)LR(0)Md(0)Ri(0)	Strawb. CR private		SignalP; Phobius	*PsAvh32* homologue; RxLR; EffectorP
11.2		g20304.t1	Contig_57	Orthogroup947	CR(14)LR(1)Md(3)Ri(3)				Antibiotic biosynthesis monooxygenase
11.2		g20307.t1	Contig_57	Orthogroup947	CR(14)LR(1)Md(3)Ri(3)				Antibiotic biosynthesis monooxygenase
11.1	Early	g7647.t1	Contig_14	Orthogroup3422	CR(14)LR(1)Md(3)Ri(3)		Species		Coil domain
11.1		g9964.t1	Contig_19	Orthogroup13882	CR(14)LR(1)Md(2)Ri(0)	*	Species	SignalP; Phobius	Coil domain
10.8		g613.t1	Contig_1	Orthogroup12809	CR(14)LR(1)Md(3)Ri(0)	*P. cactorum* private	Species	SignalP; Phobius	
10.7	Early	g4491.t1	Contig_7	Orthogroup201	CR(14)LR(1)Md(3)Ri(3)				Coil domain
10.7		g13303.t1	Contig_30	Orthogroup8651	CR(14)LR(1)Md(3)Ri(3)		Pathotype, species	SignalP; Phobius; TM domain	EffectorP
10.5		g5134.t1	Contig_8	Orthogroup6597	CR(14)LR(1)Md(3)Ri(3)		Pathotype, species	SignalP; Phobius; TM domain	RxLR
10.3		g10084.t1	Contig_20	Orthogroup1931	CR(14)LR(1)Md(3)Ri(3)			SignalP; Phobius	EffectorP
10.3	Early	g15153.t1	Contig_36	Orthogroup9085	CR(14)LR(1)Md(3)Ri(3)		spec ies	SignalP; Phobius	Protease inhibitor (Kazal-type)
10.3		g15397.t1	Contig_37	Orthogroup1297	CR(14)LR(1)Md(3)Ri(3)		pat hotype, population	SignalP; Phobius	Unamed family (PTHR34737:SF2)
10.3		g17714.t1	Contig_46	Orthogroup9681	CR(14)LR(1)Md(3)Ri(3)			SignalP; Phobius	RxLR

Transcripts were ranked by log fold change (LFC) in descending order. Gene families expanded, private to Phytophthora cactorum, or private to the strawberry crown rot (CR) lineage were identified. Functional annotation of proteins provide evidence for effector candidates.

**denotes orthogroups that could not be tested.*

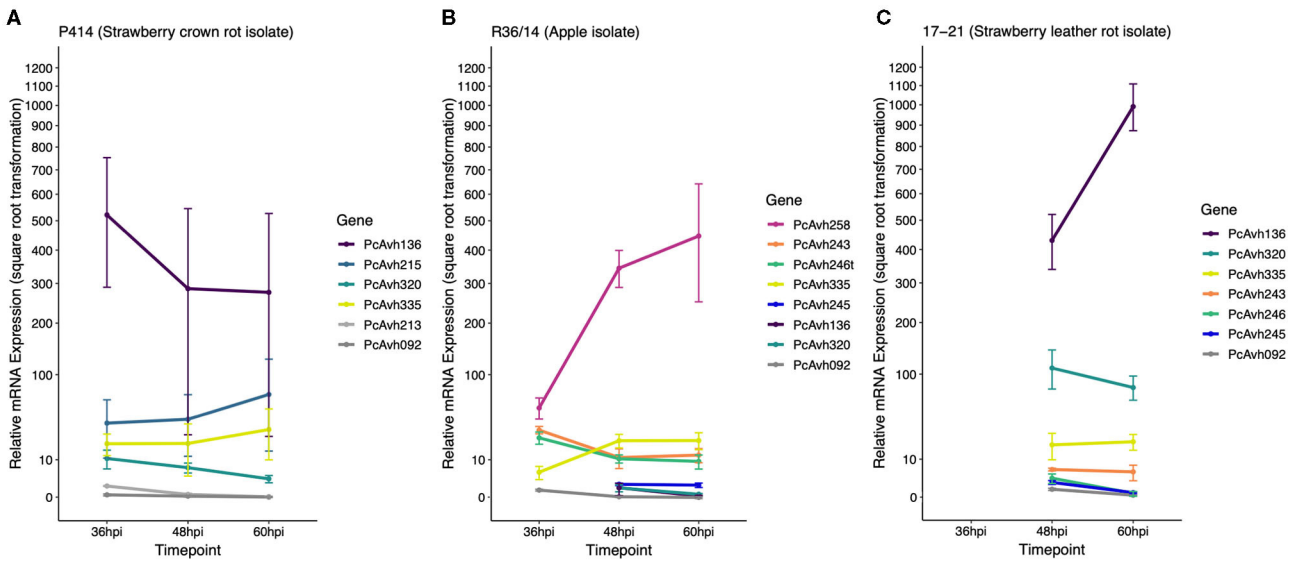


FIGURE 6 | P414, R36/14, and 17–21 RxLR genes show differential expression upon infection of strawberry fruit. **(A–C)** Relative expression of a selection of candidate RxLR genes in three representative *Phytophthora cactorum* isolates. The same gene from different genomes are indicated by the same colour. Genes of interest were normalised to two endogenous reference genes, a ribosomal 40S protein and a protein of the BAR-domain family, Pc_WS41 (Yan and Liou, 2006) and were plotted relative to the expression of the gene of interest in mycelia. Data are the mean of three biological replicates ± se. **(A)** *P. cactorum* isolate P414, strawberry crown rot pathotype representative, **(B)** *P. cactorum* isolate R36/14, apple pathotype representative, **(C)** *P. cactorum* isolate 17–21, strawberry leather rot pathotype representative.

82% pairwise aa identity, respectively, downstream of the signal peptide (26-147 aa). *Pcac1_g16973/5*, hereafter *PcAvh136*, was a gene duplication event. *Pcac1_g5134* was the only one out of eight RxLRs with a unique polymorphism (non-synonymous SNP, G53A) between strawberry CR isolates and the three apple and 17-21 isolates, that was expressed.

Looking further afield to the remaining nine RxLRs in the top 100 genes, there was another gene duplication event, *Pcac1_g16030/1*, hereafter *PcAvh320*. *PcAvh320* was identified to be a homologue to *PiAvrblb2* (Oh et al., 2009). This homologue was identified as present in all *P. cactorum* isolates. Subsequent RT-qPCR of P414 supported the RNAseq data and showed expression of the gene upon infection of strawberry fruit (**Figure 6A**). *Pcac1_g15770* (hereafter known as *PcAvh266/7/8*) was also identified to be a homologue to another known Avr gene, *PiAvrSmira1* (Rietman et al., 2012).

Unique Strawberry CR Putative Secreted Proteins Are Not All Expressed During Strawberry Infection

Of the three putative secreted proteins expanded in strawberry CR isolates, only RxLR candidate *PcAvh215* was found to be differentially expressed by P414 in the time points investigated in strawberry *in vitro* plants (with peak FPKM values of 3,364 and 4,375 in 'Emily' and 'Fenella,' respectively), and strawberry fruit (**Figure 6A**). The other RxLR candidate, *Pcac1_g22827* (hereafter *PcAvh213*), showed low levels of constitutive expression in mycelium with a peak FPKM value of 6 but was not expressed at any time point *in planta* in the RNA-Seq experiment nor subsequent *in planta* RT-qPCR experiment (**Figure 6A**). The uncharacterised secreted protein (*Pcac1_g6287*) had a peak FPKM value of 20 *in planta* but was not investigated in the subsequent RT-qPCR.

Candidate RxLR Genes and Homologues to Known Avirulence RxLR Genes Show Differential Expression Between Isolates

RT-qPCR analysis of a selection of apple lineage specific (Branch B_0; **Figure 4**) RxLR candidates show they are expressed by the apple pathotype representative R36/14 upon infection of strawberry fruit (**Figure 6B**), but not the strawberry LR representative isolate 17–21 (**Figure 6C**). The truncated *PcAvh246t* in R36/14 was expressed (**Figure 6B**) but the full-length gene, *PcAvh246*, in 17–21 was not expressed very highly (**Figure 6C**). *PC123_g25079* (hereafter *PcAvh243*) had a similar pattern of expression in R36/14 to *PcAvh246t* but was more highly expressed than *PcAvh246* in 17–21. *PC123_g17462* (hereafter *PcAvh245*) was not expressed very highly in either R36/14 or 17–21 at the time points investigated. *PC123_g15654* was not investigated by RT-qPCR. *PcAvh258*, the only RxLR candidate unique to apple isolates, was the highest expressed RxLR investigated in isolate R36/14 on strawberry fruit (**Figure 6B**).

The expression profiles of a selection of known *Avr* genes were investigated during strawberry fruit infection by the representative isolates. A homologue to *PiAvramr1* (Lin et al., 2020; *PcAvh335*) was universally expressed in the three isolates investigated (**Figure 6**, **Supplementary Table 5**). In contrast, a homologue to *PiAvramr3* (Lin et al., 2019; *PcAvh092*) was not expressed by any of the isolates at the time points

investigated (**Figure 6**, **Supplementary Table 5**). A homologue to *PiAvrblb2* (Oh et al., 2009; *PcAvh320*) was shown to have the greatest upregulation in the strawberry-infecting isolates P414 and 17–21, with lower upregulation in R36/14 (**Figure 6**, **Supplementary Table 5**). P414 was predicted to have two copies of the *PiAvrblb2* homologue adjacent to each other on contig 39. Other homologues identified but not investigated with RT-qPCR included *PiAvrvnt1* (Pel, 2010; *PcAvh428/9*) which was not expressed by P414 in the RNAseq data and *PiAvrSmira1* (Rietman et al., 2012; Pc366/7/8) which was found to be expressed *in planta* from the RNAseq data. *PcAvh258*, which was only found in apple isolates, was found to be a homologue to *PiAvr3a*. Other homologues not investigated by RT-qPCR are detailed in **Supplementary Table 5**.

The RNAseq data from strawberry showed candidate *PcAvh136* was highly expressed by P414 *in planta* and was investigated further in the three representative isolates, where it was found to be the highest expressed RxLR effector investigated in both P414 and 17-21 in strawberry fruit by RT-qPCR. The gene although present in the R36/14 genome was not expressed by the apple isolate, and therefore appears to only be expressed by the two strawberry infecting isolates. In addition, candidate *PcAvh320* appears to be expressed higher by P414 and 17-21, compared to R36/14.

DISCUSSION

Understanding the pathogenicity of plant pathogens is necessary for designing and implementing durable resistance strategies. Pathogen-host interactions are notoriously dynamic and *Phytophthora* spp. exhibit rapid adaptability to host immunity (Wang and Jiao, 2019; Chepsergon et al., 2020). *P. cactorum* is a continuing threat to strawberry and apple production, as well as more generally to forest trees and woody perennials. Our study revealed that isolates of *P. cactorum* have clear differences in pathogenicity and are generally specialised to different hosts. Sequencing of the contemporary isolates of strawberry CR revealed that there is clear separation between the *P. cactorum* isolates infecting strawberry crowns and apple. We determined that although isolates of *P. cactorum* causing strawberry CR are genetically similar globally, variation in virulence between the isolates was observed, especially on host cultivars with a higher level of resistance. Furthermore, our results highlight several RxLRs that warrant further investigation as host specificity determinants for strawberry and apple.

P. cactorum is often described as a generalist pathogen with a wide host range (Grenville-Briggs et al., 2017; Yang et al., 2018). However, our results support previous reports (Seemüller and Schmidle, 1979) which showed that *P. cactorum* isolates originating from strawberry crowns are specialised and less pathogenic on apple bark tissue than isolates originating from strawberry fruit or apple. Conversely Seemüller and Schmidle (1979) showed that apple and strawberry fruit isolates were less pathogenic on strawberry crowns. Our data showing the segregation of the strawberry CR and apple isolates into different clades demonstrates clear patterns of separation, which are backed up by analysis of population structure, potentially indicating a species complex and not a single *P. cactorum* species. Of note, the two strawberry LR isolates from our study appear to have a broader host range than either strawberry CR isolates or apple isolates and are able to cause disease in both strawberry crowns and apple excised shoots but to a more limited extent than most apple isolates. Interestingly, these isolates fall both within the clonal strawberry CR lineages and the more diverse "apple" clade. This may be what has led to confusion in the past, as it is clear that some isolates have a slightly broader host range, but potentially at the expense of enhanced virulence on a single host. However, from our work it is hard to draw conclusions from the strawberry LR isolates as only two isolates were investigated in this study. Further investigation of the isolates sequenced in this study through further RNAseq analysis of infection and the collection of additional isolates may provide a greater understanding of the components required to be a successful pathogen on both strawberry and apple. Screening all isolates on additional potential hosts was not part of the scope of this work but would be an appropriate next step to determine the host range of the isolates.

Clonal pathogen populations are a threat to global food security, as the advantageous stabilising of favourable multi-locus associations results in the rapid spread of specialised lineages within susceptible host populations. These populations can readily adapt to new introduced resistance genes, despite the proportionately moderate levels of standing genetic diversity (Stukenbrock and Bataillon, 2012), presumably due to their high census population sizes (Barton, 2010). In asexual lineages of plant pathogens such as *P. cactorum*, very little is known about the degree to which genetic variation within these lineages influences the virulence profile of populations. We observed a lack of SNP diversity between the strawberry CR isolates, despite isolates being collected from multiple countries, across multiple years, indicating clonality and the possibility of a recent bottleneck. Despite the lack of genetic diversity, substantial variation in pathogenicity on 'Elsanta' and 'Fenella' strawberry crowns was observed between the strawberry CR isolates. Further investigation is needed to understand whether this variation is either the result of what little genetic variation there is, or whether there is also stably inherited epigenetic silencing of effectors, which has recently been reported in *P. sojae* (Wang et al., 2019) or DNA methylation. N6-methylation (6 mA) was profiled in both *P. infestans* and *P. sojae*, however, its exact role in the regulation of gene expression remains to be fully examined (Chen et al., 2018). In the distantly related *P. fragariae*, another pathogen of strawberry, our previous work highlighted that variation in virulence was not related to DNA sequence variation but rather differences in the expression of putative effectors was associated with race structure (Adams et al., 2020). It was proposed that silencing of *Avr* RxLR effectors enables isolates to evade recognition in plants possessing the corresponding *R* gene. The exact mechanism of silencing was not determined during the study, but we postulated that variation was attributed to either control in *trans* or stable forms of epigenetic modifications were regulating gene expression (Adams et al., 2020). Little is known about how filamentous pathogens adapt to host plants by factors other than DNA sequence polymorphism. This study revealed that over half (86) of our effector candidates were not expressed or showed

low-expression *in planta*. This indicates that there is pervasive silencing of effectors, highlighting the crucial role of RNAseq data when attempting to determine effector candidates for onward study.

Non-host resistance (NHR), described roughly as the ability of a plant species to ward off the colonisation of all genotypes of a pathogen species, is poorly understood but often considered the most durable form of resistance, due to the large number of independent resistance mechanisms that are likely acting. Studies in a wide range of pathogens provide evidence for the role of secreted effectors in determining host range (Lee et al., 2017). A classic example is the deletion of a single effector gene, *hopQ1-1* in *Pseudomonas syringae* pv. *tomato*, enabling the pathogen to extend its host range within the Solenaceae to the non-host plant species, *Nicotiana benthamiana* (Wei et al., 2007). However, non-host resistance ranges from single effectors controlling host range to potentially many tens of effectors. A study of 54 *P. infestans* RxLR effectors in the non-host pepper uncovered that up to 36 could be recognised and led to hypersensitive response in some accessions, suggesting that the recognition of multiple *P. infestans* effectors leads to NHR (Lee et al., 2014). This large number of recognised effectors may be due to the evolutionary differences between wild potato and *Capsicum* species – reported to have split around 20 Ma (Särkinen et al., 2013), which although geographically overlapping have distinct ecological niches (Rumold and Aldenderfer, 2016; Barboza et al., 2019). However, it is possible that *Phytophthora* isolates may carry within their genome many more effectors than they actually express (as we highlight above) and so what matters is which effectors are expressed during infection. What also matters is not just the antigenic potential of the effectors that are expressed, but the function of the effector network in the infection process. It is well-known that some effectors are able to mask the avirulence potential of other effectors and therefore evidence for activation of the hypersensitive response alone is also insufficient to assess the functional consequence of a lineage possessing any single effector (Derevnina et al., 2021).

From our work it is not possible to tell how many effectors are causing the observed differences in host range between the clonal CR lineage and isolates infective on apple. The reciprocally incompatible interactions that we observe could be due to effectors either being recognised or the inability of the lineage to efficiently manipulate the host. However, our work did identify candidate RxLR's which may be good candidates for host specificity determinants on strawberry and apple. Two candidate host-range determinants on strawberry crowns and apple that warrant further investigation are PcAvh215 and PcAvh258, respectively. The strawberry LR isolate, 17–21 does not possess PcAvh215 or PcAvh258 but is pathogenic on both strawberry and apple, to a limited extent compared to isolates recovered from the specific host tissue. This discounts these effectors from being solely responsible for pathogenicity on either host, but they possibly enable the isolates possessing them to have greater virulence on the respective host. Candidate effector *PcAvh136* was found to be expressed by both the strawberry CR isolate P414 and the strawberry LR isolate 17–21, but not the apple isolate R36/14 upon infection of strawberry fruit, indicating it may possibly be a determinant for pathogenicity on strawberry. It may also be that *PcAvh215* enables isolates possessing it to be more pathogenic on strawberry crowns (all strawberry CR isolates and strawberry LR isolate 11–40), but further RNAseq and transformation of it into non-PcAvh215 containing isolates such as 17–21 would be required to investigate this hypothesis further. Of the effectors that were found to be polymorphic between apple and strawberry only one showed evidence for expression in strawberry.

Understanding effector profiles has aided the characterisation of resistance genes (Armstrong et al., 2005; Poppel et al., 2008; Champouret et al., 2009; Oh et al., 2009; Gilroy et al., 2011; Rietman et al., 2012; Sugimoto et al., 2012), and thereby provided an insight into resistance durability in the field. Homologues to major gene targets (homologues to known *Avr* genes) that are expressed by *P. cactorum* include, *PiAvr1* (Champouret et al., 2009) *PiAvrblb1* (Champouret et al., 2009), *PiAvrblb2* (Oh et al., 2009), *PiAvrSmira1* (Rietman et al., 2012) and *PiAvramr1* (Lin et al., 2020). It would be interesting to investigate if homologues to the resistance genes *Rpiblb1*, *Rpiblb2*, and *RpiSmira1* are involved in resistance to *P. cactorum* in strawberry and apple. In addition, there are multiple major gene resistance targets in *P. cactorum* that are present but were not expressed in the *in planta* RNAseq time course of the strawberry CR isolate P414, for example, *PiAvrvnt1* (Pel, 2010), *PiAvr4* (Poppel et al., 2008), *PiAvr8/PiAvrSmira2* (Rietman et al., 2012), and *PiAvramr3* (Lin et al., 2019), it is unknown if these play a role in virulence in *P. cactorum*.

Copy number variation (CNV) is also another form of genetic adaptation that has been identified in *Phytophthora* genomes (Qutob et al., 2009). For example, *P. sojae, P. parasitica*, and *P. infestans* genomes encode at least 2, 8, and 11 *Avrblb2* homologues, respectively (Naveed et al., 2019). In the most contiguous *P. cactorum* genome of strawberry CR isolate P414, two homologues, identical to each other (*PcAvh320*) were predicted adjacent to each other in the genome and were expressed upon infection of strawberry roots. This example of effector duplication therefore supplies the raw materials for adaptive evolution of the gene into novel functions in *P. cactorum*. Further long-read sequencing of additional isolates is required to determine the extent of CNV in *P. cactorum* as it is clear that short-read sequencing fails to detect instances of gene duplication with the same fidelity as long-read sequencing.

In planta transcriptome sequencing of further isolates will be able to investigate both effector expression variation and the role for CNV's, alongside functional characterisation and genetic analysis of the basis of resistance to *Phytophthora* in both strawberry and apple. This is non-trivial in both of these crops due to the lack of appropriate methods to transiently induce expression, as infiltration techniques into leaves are challenging due to leaf properties and until extremely recently genome sequences of the octoploid strawberry were not available to aid with the characterisation of putative resistance genes (Edger et al., 2019).

As with other *Phytophthora* spp. genomes, the advent of long-read sequencing technology has improved the generation of genome assemblies (Adams et al., 2020; Shi et al., 2021).

In our study, a single strawberry CR isolate, P414, was sequenced with long read technology and was found to possess a greater genome size, number of gene models and greater numbers of genes, including RxLR and CRN effectors, compared to the other strawberry CR isolates sequenced by Illumina short-read technology. However, all comparative analyses (phylogenetic, orthogroup, expansion/contraction, and variant calling) performed within this study focus on shared characters across 13 isolates, comprised of this single long-read assembly and 12 short-read assemblies, and are not impacted by the prediction of additional genomic content. Therefore, the identification of potential host-determinants is a conservative one and there is the possibility of additional host-adapted effectors to be discovered, as additional long-read data is generated.

CONCLUSION

This study provides further evidence that *P. cactorum* should be regarded as a species complex and not a single species, as it comprises of distinct phylogenetic lineages that resolve groups of isolates with distinct effector profiles and displaying host-preference. Pathotype specific effector genes, such as homologues to *PsAvh32* (*PcAvh215*) and *PiAvr3a* (*PcAvh258*) may play roles in specialisation of *P. cactorum* to strawberry and apple, respectively. However, functional analysis is required to validate these genes as determinants of pathogenicity in their respective host.

Further questions that also remain unanswered are to what extent do the expression profiles of effectors in different isolates affect pathogenicity? This highlights the need for further RNAseq from multiple isolates, as well as knockouts to disentangle which effectors are differential for pathogenicity in strawberry and apple? This will help us understand what are the key processes that underpin variation in virulence in *P. cactorum*.

This study raises questions about the strategy for effector-informed breeding. We know that clonal lineages predominate in strawberry, as has often been described in other agricultural-associated pathosystems (Hessenauer et al., 2021). We have shown that there are highly expressed lineage-specific effectors within the clonal lineage of CR along with CNV in highly expressed effectors. We hypothesise that these are associated with increased virulence and potentially also pathogenicity itself. If these were then used to screen for and then deploy resistance against them, then the question arises as to what would the response be, if the pathogen were to adapt? It could be that there are many other effectors that could play a similar role, that are currently silent within the genome, or conversely that through silencing of these effectors, host resistance may be evaded, but at a fitness cost to the pathogen. However, this remains to be seen. Conversely, we see that many "core" effectors that are conserved across lineages on different hosts are present but not expressed, so these, while we may assume they are important from DNA sequence data alone, they are clearly dispensable for pathogenicity, indicating that the idea of core effectors needs to expand beyond simply an analysis of their presence within a genome. Therefore, we must find ways to understand how the network of effectors functions in any given host and whether there are critical effector combinations that if targeted by multiple resistance genes would lead to a durable resistance.

DATA AVAILABILITY STATEMENT

The datasets, including raw data, assemblies, and annotations, generated for this study are available on NCBI GenBank as part of BioProjects PRJNA383548 and PRJNA391273, accession numbers are shown in **Table 1**. The BioSample IDs for the RNAseq data are SAMN18192675-SAMN18192689.

AUTHOR CONTRIBUTIONS

RH, CN, and AA devised the study, conceived, and drafted the manuscript. CN performed the experimental work, with input from LL, HB, and ML. AA performed the bioinformatic analyses including genome assembly, annotation, orthology gene expression, and variant calling analyses with input from MS. All authors read and approved the submission.

FUNDING

This research was supported by grants awarded to RH from the Biotechnology and Biological Sciences Research Council (BBSRC–BB/K017071/1, BB/K017071/2, and BB/N006682/1) and ML who is receipt of Ph.D., funding from the BBSRC Collaborative Training Partnership for Crop Research (CTP_FCR2018-5, AHDB Grant ST/TF170).

ACKNOWLEDGMENTS

The authors gratefully acknowledge the NIAB EMR Plant Clinic, Prof. May Bente Brurberg, Drs Thijs van Dijk, Vance M. Whitaker, and Prof. Natalia A. Peres for providing *P. cactorum* isolates. The authors also thank Dr Thomas M. Adams for useful discussions and the East Malling Strawberry Breeding Club for access to strawberry material. In addition, we wish to thank Adam B. Whitehouse, Andrew J. Passey, César Marina-Montes, and Joseph Hutchings for the help received in the preparation and setting up of experiments. Work was carried out under the terms of DEFRA Plant Health Licence 6996/221427 held by RH/CN.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.679936/full#supplementary-material

Supplementary Figure 1 | Example experimental setup of *in vitro Fragaria* × *ananassa* plants. **(A)** Strawberry plants positioned in petri dishes, on top of agar, **(B)** aluminium foil casing and upright positioning of plates in the growth incubator.

Supplementary Figure 2 | All *Phytophthora cactorum* isolates tested were able to cause disease in strawberry fruit. Percentage of symptomatic strawberry

'Elsanta' fruit after artificial inoculation *Phytophthora cactorum* and *Phytophthora idaei* zoospores, from three separate experiments.

Supplementary Figure 3 | Agarose gel electrophoresis of RT-PCR β-tubulin reactions on representative samples from inoculation time course experiment of *Phytophthora cactorum* isolate P414 on the 'Emily' cultivar of *Fragaria* × *ananassa*. L: 100 bp DNA Ladder (New England Biolabs). 1: Mock inoculated 'Emily.' 2: 6 h post inoculation (hpi). 3: 12 hpi. 4: 24 hpi. 5: 48 hpi. 6: 72 hpi. 7: 96 hpi. 8: 120 hpi. 9: gDNA from 'Emily' (negative control). 10: gDNA from P414 mycelium (positive control). 11: dH_2O (negative control).

Supplementary Table 1 | Characteristics of primers used in this study.

Supplementary Table 2 | Details of genes expanded and contracted in the *Phytophthora cactorum* strawberry crown rot and apple lineages.

Supplementary Table 3 | Summary of identified *Phytophthora cactorum* Avirulence gene homologues.

Supplementary Table 4 | *Phytophthora cactorum* isolate P414 annotation table and *in planta* RNAseq analysis from *Fragaria* × *ananassa*.

Supplementary Table 5 | Summary of homologues to known avirulence RxLR genes in *Phytophthora cactorum*. The region of genes for comparison are detailed in column three, noted as amino acid (AA) sequence after signal peptide (SP).

REFERENCES

Adams, T. M., Armitage, A. D., Sobczyk, M. K., Bates, H. J., Tabima, J. F., Kronmiller, B. A., et al. (2020). Genomic investigation of the strawberry pathogen Phytophthora fragariae indicates pathogenicity is associated with transcriptional variation in three key races. *Front. Microbiol.* 11:490. doi: 10.3389/fmicb.2020.00490

Alexander, B., and Stewart, A. (2001). Glasshouse screening for biological control agents of *Phytophthora cactorum* on apple (*Malus domestica*). *N. Zeal. J. Crop Horticult. Sci.* 29, 159–169. doi: 10.1080/01140671.2001.9514174

Anderson, R. G., Deb, D., Fedkenheuer, K., and McDowell, J. M. (2015). Recent progress in RXLR effector research. *Mol. Plant Microbe Interact.* 28, 1063–1072. doi: 10.1094/mpmi-01-15-0022-cr

Armitage, A. D., Lysøe, E., Nellist, C. F., Lewis, L. A., Cano, L. M., Harrison, R. J., et al. (2018). Bioinformatic characterisation of the effector repertoire of the strawberry pathogen Phytophthora cactorum. *PLoS ONE* 13:e0202305. doi: 10.1371/journal.pone.0202305

Armstrong, M. R., Whisson, S. C., Pritchard, L., Bos, J. I. B., Venter, E., Avrova, A. O., et al. (2005). An ancestral oomycete locus contains late blight avirulence gene Avr3a, encoding a protein that is recognized in the host cytoplasm. *Proc. Natl. Acad. Sci. U.S.A.* 7766–7771. doi: 10.1073/pnas.0500113102

Aronesty, E. (2013). Comparison of sequencing utility programs. *Open Bioinform. J.* 7, 1–8 doi: 10.2174/1875036201307010001

Auwera, G. A. V., Carneiro, M. O., Hartl, C., Poplin, R., Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinform.* 43, 1–33. doi: 10.1002/0471250953.bi1110s43

Ball, T. B., Plummer, F. A., and HayGlass, K. T. (2003). Improved mRNA quantitation in LightCycler RT-PCR. *Int. Arch. Allergy Immunol.* 130, 82–86. doi: 10.1159/000068372

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Barboza, G. E., Carrizo García, C., Leiva González, S., Scaldaferro, M., and Reyes, X. (2019). Four new species of capsicum (*Solanaceae*) from the tropical andes and an update on the phylogeny of the genus. *PLoS ONE* 14:e0209792. doi: 10.1371/journal.pone.0209792

Barton, N. (2010). Understanding adaptation in large populations. *PLoS Genetics* 6:e1000987. doi: 10.1371/journal.pgen.1000987

Bhat, R. G., Colowit, P. M., Tai, T. H., Aradhya, M. K., and Browne, G. T. (2006). Genetic and pathogenic variation in phytophthora cactorum affecting fruit and nut crops in California. *Plant Dis.* 90, 161–169. doi: 10.1094/pd-90-0161

Boyer, L. R., Feng, W., Gulbis, N., Hajdu, K., Harrison, R. J., Jeffries, P., et al. (2016). The use of arbuscular mycorrhizal fungi to improve strawberry production in coir substrate. *Front. Plant Sci.* 7, 83–89. doi: 10.3389/fpls.2016.01237

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348

Champouret, N., Bouwmeester, K., Rietman, H., Lee, T., Maliepaard, C., Heupink, A., et al. (2009). Phytophthora infestans isolates lacking class I ipiO variants are virulent on Rpi-blb1 potato. *Mol. Plant Microbe Interact.* 22, 1535–1545. doi: 10.1094/MPMI-22-12-1535

Chen, H., Shu, H., Wang, L., Zhang, F., Li, X., Ochola, S. O., et al. (2018). Phytophthora methylomes are modulated by 6mA methyltransferases and associated with adaptive genome regions. *Genome Biol.* 19:181. doi: 10.1186/s13059-018-1564-4

Chepsergon, J., Motaung, T. E., Bellieny-Rabelo, D., and Moleleki, L. N. (2020). Organize, don't agonize: strategic success of phytophthora species. *Microorganisms* 8:917. doi: 10.3390/microorganisms8060917

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* 6, 80–92. doi: 10.4161/fly.19695

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43, 491–498. doi: 10.1038/ng.806

Derevnina, L., Contreras, M. P., Adachi, H., Cruces, A. V., Xie, R., Sklenar, J., et al. (2021). Plant pathogens convergently evolved to counteract redundant nodes of an NLR immune receptor network. *bioRxiv*. 1–15. doi: 10.1101/2021.02.03.429184

Deutschmann, V. F. (1954). Eine wurzelfäule an erdbeeren, hervorgerufen durch *Phytophthora cactorum* (Leb. et Cohn) Schröt. *Nachr. Deutsch. Pflanzenschutzd.* 6, 7–9

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21 doi: 10.1093/bioinformatics/bts635

Edger, P. P., Poorten, T. J., VanBuren, R., Hardigan, M. A., Colle, C., McKain, M. R., et al. (2019). Origin and evolution of the octoploid strawberry genome. *Nat. Genet.* 51, 541–547. doi: 10.1038/s41588-019-0356-4

Ellis, M. A., and Grove, G. G. (1983). Leather rot in ohio strawberries. *Plant Dis.* 67:549.

Erwin, D. C., and Ribeiro, O. K. (1996). *Phytophthora Diseases Worldwide*. St. Paul, MN: American Phytopathological Society Press.

Gilroy, E. M., Taylor, R. M., Hein, I., Boevink, P., Sadanandom, A., and Birch, P. R. J. (2011). CMPG1-dependent cell death follows perception of diverse pathogen elicitors at the host plasma membrane and is suppressed by phytophthora infestans RXLR effector AVR3a. *New Phytol.* 190, 653–666. doi: 10.1111/j.1469-8137.2011.03643.x

Grenville-Briggs, L. J., Kushwaha, S. K., Cleary, M. R., Witzell, J., Savenkov, E. I., Whisson, S. C., et al. (2017). Draft genome of the oomycete pathogen Phytophthora cactorum strain LV007 isolated from European beech (*Fagus sylvatica*). *Genom. Data* 12, 155–156. doi: 10.1016/j.gdata.2017.05.010

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., et al. (2009). Genome sequence and analysis of the irish potato famine pathogen *Phytophthora infestans*. *Nature* 461, 393–398. doi: 10.1038/nature08358

Hantula, J., Lilja, A., Nuorteva, H., Parikka, P., and Werres, S. (2000). Pathogenicity, morphology and genetic variation of *Phytophthora cactorum* from strawberry, apple, rhododendron, and silver birch. *Mycol. Res.* 104, 1062–1068. doi: 10.1017/S0953756200002999

Hantula, J., Lilja, A., and Parikka, P. (1997). Genetic variation and host specificity of *Phytophthora cactorum* isolated in Europe. *Mycol. Res.* 101, 565–572.

Harris, D. C. (1991). The phytophthora diseases of apple. *J. Horticult. Sci.* 66, 513–544. doi: 10.1080/00221589.1991.11516181

Hessenauer, P., Feau, N., Gill, U., Schweddinger, B., Brar, G. S., and Hamelin, R. C. (2021). Evolution and adaptation of forest and crop pathogens in the anthropocene. *Phytopathology* 111, 49–67. doi: 10.1094/PHYTP-08-20-0358-FI

Kamoun, S., West, P., Jong, A. J., Groot, K. E., Vleeshouwers, V. G. A. A., and Govers, F. (1997). A gene encoding a protein elicitor of phytophthora infestans is down-regulated during infection of potato. *Mol. Plant Microbe Interact.* 10, 13–20. doi: 10.1094/mpmi.1997.10.1.13

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evolut.* 30, 772–780. doi: 10.1093/molbev/mst010

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate longread assembly via adaptive k-mer weighting and repeating. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Lee, H., Kim, S., Oh, S., Yeom, S., Kim, S., Kim, M., et al. (2014). Multiple recognition of RXLR effectors is associated with nonhost resistance of pepper against *Phytophthora infestans*. *New Phytol.* 203, 926–938. doi: 10.1111/nph.12861

Lee, H.-A., Lee, H.-Y., Seo, E., Lee, J., Kim, S.-B., Oh, S., et al. (2017). Current understandings of plant nonhost resistance. *Mol. Plant Microbe Interact.* 30, 5–15. doi: 10.1094/mpmi-10-16-0213-cr

Lilja, A., Karjalainen, R., Parikka, P., Kammiovirta, K., and Nuorteva, H. (1998). Pathogenicity and genetic variation of Phytophthora cactorum from silver birch and strawberry. *Euro. J. Plant Pathol.* 104, 529–535. doi: 10.1023/A:1008644804415

Lin, X., Song, T., Fairhead, S., Witek, K., Jouet, A., Jupe, F., et al. (2020). Identification of Avramr1 from *Phytophthora infestans* using long read and cDNA pathogen-enrichment sequencing (PenSeq). *Mol. Plant Pathol.* 21, 1502–1512. doi: 10.1111/mpp.12987

Lin, X., Witek, K., Witek, A., McLellan, H., Nellist, C. F., Armitage, A. D., et al. (2019). The recognition of conserved RxLR effectors of *Phytophthora species* might help to defeat multiple oomycete diseases. *Mol. Plant Microbe Interact.* 32, S11–S1212. doi: 10.1094/mpmi-32-10-s1.1

Liu, K., Linder, C. R., and Warnow, T. (2011). RAxML and FastTree: comparing two methods for large-scale maximum likelihood phylogeny estimation. *PLoS ONE* 6:e27731. doi: 10.1371/journal.pone.0027731

Longmuir, A. L., Beech, P. L., and Richardson, M. F. (2018). Draft genomes of two Australian strains of the plant pathogen, *Phytophthora cinnamomi*. *F1000Research* 6, 1972. doi: 10.12688/f1000research.12867.1

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 1–21. doi: 10.1186/s13059-014-0550-8

Luberti, M., Litthauer, S., Dunwell, J. M., Fernández Fernández, F., and Nellist, C. F. (2021). Response of apple (*Malus domestica*) accessions to UK *Phytophthora cactorum* isolates in cut-shoot tests. *Acta Horticult.* 1307, 369–374. doi: 10.17660/ActaHortic.2021.1307.56

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

Naveed, Z. A., Bibi, S., and Ali, G. S. (2019). The phytophthora RXLR effector Avrblb2 modulates plant immunity by interfering with Ca2+ signaling pathway. *Front. Plant Sci.* 10:374. doi: 10.3389/fpls.2019.00374

Nellist, C. F., Vickerstaff, R. J., Sobczyk, M. K., Marina-Montes, C., Wilson, F. M., Simpson, D. W., et al. (2019). Quantitative trait loci controlling Phytophthora cactorum resistance in the cultivated octoploid strawberry (fragaria x ananassa). *Horticult. Res.* 6:60. doi: 10.1038/s41438-019-0136-4

Oh, S. K., Young, C., Lee, M., Oliva, R., Bozkurt, T. O., Cano, L. M., et al. (2009). In planta expression screens of phytophthora infestans RXLR effectors reveal diverse phenotypes, including activation of the solanum bulbocastanum disease resistance protein Rpi-blb2. *Plant Cell* 21, 2928–2947. doi: 10.1105/tpc.109.068247

Orsomando, G., Lorenzi, M., Raffaelli, N., Rizza, M. D., Mezzetti, B., and Ruggieri, S. (2001). Phytotoxic protein PcF, purification, characterization, and cDNA sequencing of a novel hydroxyproline-containing factor secreted by the strawberry pathogen *Phytophthora cactorum*. *J. Biol. Chem.* 276, 21578–21584. doi: 10.1074/jbc.m101377200

Patro, R., Duggal, G., Love, M. I., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419. doi: 10.1038/nmeth.4197

Pel, M. A. (2010). *Mapping, isolation and characterization of genes responsible for late blight resistance in potato* (Ph.D. thesis), Wageningen, UR, 210. Available online at: https://edepot.wur.nl/138132

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res.* 29:e45. doi: 10.1093/nar/29.9.e45

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., and Lercher, M. J. (2014). PopGenome: an efficient swiss army knife for population genomic analyses in R. *Mol. Biol. Evolut.* 31, 1929–1936. doi: 10.1093/molbev/msu136

Poppel, P. M. J. A., Guo, J., Vondervoort, P. J. I., Jung, M. W. M., Birch, P. R. J., Whisson, S. C., et al. (2008). The Phytophthora infestans avirulence gene Avr4 encodes an RXLR-dEER effector. *Mol. Plant Microbe Interact.* 21, 1460–1470. doi: 10.1094/MPMI-21-11-1460

Qutob, D., Tedman-Jones, J., Dong, S., Kuflu, K., Pham, H., Wang, Y., et al. (2009). Copy number variation and transcriptional polymorphisms of phytophthora sojae RXLR effector genes Avr1a and Avr3a. *PLoS ONE* 4:e5066. doi: 10.1371/journal.pone.0005066

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Available online at: https://www.R-project.org/

Raj, A., Stephens, M., and Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 197, 573–589. doi: 10.1534/genetics.114.164350

Rietman, H., Bijsterbosch, G., Cano, L. M., Lee, H.-R., Vossen, J. H., Jacobsen, E., et al. (2012). Qualitative and quantitative late blight resistance in the potato cultivar sarpo mira is determined by the perception of five distinct RXLR effectors. *Mol. Plant Microbe Interact.* 25, 910–919. doi: 10.1094/MPMI-01-12-0010-R

Rose, D. H. (1924). Leather rot of strawberries. *J. Agricult. Sci.* 28, 357–376.

Rumold, C. U., and Aldenderfer, M. S. (2016). Late archaic-early formative period microbotanical evidence for potato at jiskairumoko in the titicaca basin of southern Peru. *Proc. Natl. Acad. Sci. U.S.A.* 113, 13672–13677. doi: 10.1073/pnas.1604265113

Särkinen, T., Bohs, L., Olmstead, R. G., and Knapp, S. (2013). A phylogenetic framework for evolutionary study of the nightshades (*Solanaceae*): a dated 1000-tip tree. *BMC Evolut. Biol.* 13:214. doi: 10.1186/1471-2148-13-214

Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27, 863–864. doi: 10.1093/bioinformatics/btr026

Schulze-Lefert, P., and Panstruga, R. (2011). A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends Plant Sci.* 16, 117–125. doi: 10.1016/j.tplants.2011.01.001

Seemüller, E., and Schmidle, A. (1979). Einfluβ der herkunft von phytophthora cactorum-isolaten auf ihre virulenz an apfelrinde, erdbeerrhizomen und erdbeerfrüchten. *Phytopathology* 94, 218–225. doi: 10.1111/j.1439-0434.1979.tb01553.x

Shi, J., Ye, W., Ma, D., Yin, J., Zhang, Z., Wang, Y., et al. (2021). Improved whole genome sequence of Phytophthora capsici generated by long-read sequencing. *Mol. Plant Microbe Interact.* doi: 10.1094/mpmi-12-20-0356-a. [Epub ahead of print].

Shulaev, V., Sargent, D. J., Crowhurst, R. N., Mockler, T. C., Folkerts, O., Delcher, A. L., et al. (2011). The genome of woodland strawberry (*Fragaria vesca*). *Nat. Genet.* 43, 109–116. doi: 10.1038/ng.740

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351

Stam, R., Jupe, J., Howden, A. J. M., Morris, J. A., Boevink, P. C., Hedley, P. E., et al. (2013). Identification and characterisation CRN effectors in phytophthora capsici shows modularity and functional diversity. *PLoS ONE* 8:e59517. doi: 10.1371/journal.pone.0059517.s005

Stam, R., Mantelin, S., McLellan, H., and Thilliez, G. (2014). The role of effectors in nonhost resistance to filamentous plant pathogens. *Front. Plant Sci.* 5:582. doi: 10.3389/fpls.2014.00582/abstract

Stensvand, A., Herrero, M. L., and Talg,ø, V. (1999). Crown rot caused by *Phytophthora cactorum* in norwegian strawberry production. *EPPO Bull.* 29, 155–158

Stukenbrock, E. H., and Bataillon, T. (2012). A population genomics perspective on the emergence and adaptation of new plant pathogens in agro-ecosystems. *PLoS Pathog.* 8, e1002893–e1002894. doi: 10.1371/journal.ppat.1002893

Sugimoto, T., Kato, M., Yoshida, S., Matsumoto, I., Kobayashi, T., Kaga, A., et al. (2012). Pathogenic diversity of phytophthora sojae and breeding strategies to develop phytophthora-resistant soybeans. *Breed. Sci.* 61, 511–522. doi: 10.1270/jsbbs.61.511

Taylor, A., Vágány, V., Jackson, A. C., Harrison, R. J., Rainoni, A., and Clarkson, J. P. (2016). Identification of pathogenicity-related genes in *Fusarium oxysporum* f. sp. *cepae*. *Mol. Plant Pathol.* 17, 1032–1047. doi: 10.1111/mpp.12346

Thomidis, T. (2003). Variability in pathogenicity among greek isolates of *Phytophthora cactorum* to four peach rootstocks. *Austr. J. Exp. Agricult.* 43, 99–95. doi: 10.1071/ea01203

Torto, T. A., Li, S., Styer, A., Huitema, E., Testa, A., Gow, N. A. R., et al. (2003). EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen phytophthora. *Genome Res.* 13, 1675–1685. doi: 10.1101/gr.910003

Wala, J. A., Bandopadhayay, P., Greenwald, N. F., O'Rourke, R., Sharpe, T., Stewart, C., et al. (2018). SvABA: genome-wide detection of structural variants and indels by local assembly. *Genome Res.* 28, 581–591. doi: 10.1101/gr.221028.117

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumer, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* 9:e112963. doi: 10.1371/journal.pone.0112963

Wang, L., Chen, H., Li, J., Shu, H., Zhang, X., Wang, Y., et al. (2019). Effector gene silencing mediated by histone methylation underpins host adaptation in an oomycete plant pathogen. *Nucleic Acids Res.* 313, 1261–1210. doi: 10.1093/nar/gkz1160

Wang, Q., Han, C., Ferreira, A. O., Yu, X., Ye, W., Tripathy, S., et al. (2011). Transcriptional programming and functional interactions within the *Phytophthora sojae* RXLR effector repertoire. *Plant Cell* 23, 2064–2086. doi: 10.1105/tpc.111.086082

Wang, W., and Jiao, F. (2019). Effectors of Phytophthora pathogens are powerful weapons for manipulating host immunity. *Planta* 250, 413–425. doi: 10.1007/s00425-019-03219-x

Wang, Y., and Wang, Y. (2018a). Phytophthora sojae effectors orchestrate warfare with host immunity. *Curr. Opin. Microbiol.* 46, 7–13. doi: 10.1016/j.mib.2018.01.008

Wang, Y., and Wang, Y. (2018b). Trick or treat: microbial pathogens evolved apoplastic effectors modulating plant susceptibility to infection. *Mol. Plant Microbe Interact.* 31, 6–12. doi: 10.1094/mpmi-07-17-0177-fi

Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., et al. (2017). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evolut.* 35, 543–548. doi: 10.1093/molbev/msx319

Wawra, S., Belmonte, R., Löbach, L., Saraiva, M., Willems, A., and West, P. (2012). Secretion, delivery and function of oomycete effector proteins. *Curr. Opin. Microbiol.* 15, 685–691. doi: 10.1016/j.mib.2012.10.008

Wedgwood, E., Berrie, A., Passey, T., Hall, A., and Xu, X. (2020). *Improving Integrated Disease Management in Strawberry*. Available online at: https://projectblue.blob.core.windows.net/media/Default/Horticulture/SF%20157_Report_Final_2020.pdf

Wei, C., Kvitko, B. H., Shimizu, R., Crabill, E., Alfano, J. R., Lin, N., et al. (2007). A *Pseudomonas syringae* pv. tomato DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. *Plant J.* 51, 32–46. doi: 10.1111/j.1365-313x.2007.03126.x

Wilcox, W. F., Scott, P. H., Hamm, P. B., Kennedy, D. M., Duncan, J. M., Brasier, C. M., et al. (1993). Identity of a *Phytophthora species* attacking raspberry in Europe and North America. *Mycol. Res.* 97, 817–831. doi: 10.1016/s0953-7562(09)81157-x

Win, J., Chaparro-Garcia, A., Belhaj, K., Saunders, D. G. O., Yoshida, K., Dong, S., et al. (2012). Effector biology of plant-associated organisms: concepts and perspectives. *Cold Spring Harbor Symp. Quant. Biol.* 77, 235–247. doi: 10.1101/sqb.2012.77.015933

Yan, H.-Z., and Liou, R.-F. (2006). Selection of internal control genes for real-time quantitative RT-PCR assays in the oomycete plant pathogen *Phytophthora parasitica*. *Fungus Genet. Biol.* 43, 430–438. doi: 10.1016/j.fgb.2006.01.010

Yang, M., Duan, S., Mei, X., Huang, H., Chen, W., Liu, Y., et al. (2018). The *Phytophthora cactorum* genome provides insights into the adaptation to host defense compounds and fungicides. *Sci. Rep.* 8:6534. doi: 10.1038/s41598-018-24939-2

Yang, X., Tyler, B. M., and Hong, C. (2017). An expanded phylogeny for the genus phytophthora. *IMA Fungus* 8, 355–384. doi: 10.5598/imafungus.2017.08.02.09

Yu, D., Tang, H., Zhang, Y., Du, Z., Yu, H., and Chen, Q. (2012). Comparison and improvement of different methods of RNA isolation from strawberry (Fragria x ananassa). *J. Agricult. Sci.* 4, 51–56. doi: 10.5539/jas.v4n7p51

Yu, G., Yu, G., Smith, D. K., Smith, D. K., Zhu, H., Zhu, H., et al. (2016). ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evolut.* 8, 28–36. doi: 10.1111/2041-210x.12628

Zhang, C., Rabiee, M., Sayyari, E., and Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinform.* 19:153. doi: 10.1186/s12859-018-2129-y

Zhang, S.-D., Jin, J.-J., Chen, S.-Y., Chase, M. W., Soltis, D. E., Li, H.-T., et al. (2017). Diversification of rosaceae since the late cretaceous based on plastid phylogenomics. *New Phytol.* 214, 1355–1367. doi: 10.1111/nph.14461

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Citation: Nellist CF, Armitage AD, Bates HJ, Sobczyk MK, Luberti M, Lewis LA and Harrison RJ (2021) Comparative Analysis of Host-Associated Variation in Phytophthora cactorum. Front. Microbiol. 12:679936. doi: 10.3389/fmicb.2021.679936

frontiers in Microbiology

ORIGINAL RESEARCH
published: 28 July 2021
doi: 10.3389/fmicb.2021.712701

Transcriptome Variations in *Verticillium dahliae* in Response to Two Different Inorganic Nitrogen Sources

*Chen Tang[1], Wenwen Li[1], Steven J. Klosterman[2] and Yonglin Wang[1]**

[1] Beijing Key Laboratory for Forest Pest Control, College of Forestry, Beijing Forestry University, Beijing, China, [2] Agricultural Research Service, United States Department of Agriculture, Salinas, CA, United States

OPEN ACCESS

Edited by:
Danyu Shen,
Nanjing Agricultural University, China

Reviewed by:
Xiong Zhang,
Oil Crops Research Institute, Chinese Academy of Agricultural Sciences, China
Linlin Chen,
Henan Agricultural University, China
Hongjie Feng,
State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, China

***Correspondence:**
Yonglin Wang
ylwang@bjfu.edu.cn

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *21 May 2021*
Accepted: *01 July 2021*
Published: *28 July 2021*

Citation:
Tang C, Li W, Klosterman SJ and Wang Y (2021) Transcriptome Variations in Verticillium dahliae in Response to Two Different Inorganic Nitrogen Sources. Front. Microbiol. 12:712701. doi: 10.3389/fmicb.2021.712701

The fungus *Verticillium dahliae* causes vascular wilt disease on hundreds of plant species. The main focus of the research to control this fungus has been aimed at infection processes such as penetration peg formation and effector secretion, but the ability of the fungus to acquire and utilize nutrients are often overlooked and may hold additional potential to formulate new disease control approaches. Little is known about the molecular mechanisms of nitrogen acquisition and assimilation processes in *V. dahliae*. In this present study, RNA sequencing and gene expression analysis were used to examine differentially expressed genes in response to the different nitrogen sources, nitrate and ammonium, in *V. dahliae*. A total of 3244 and 2528 differentially expressed genes were identified in response to nitrate and ammonium treatments, respectively. The data indicated nitrate metabolism requires additional energy input while ammonium metabolism is accompanied by reductions in particular cellular processes. Gene ontology and Kyoto Encyclopedia of Genes and Genomes pathway analyses of DEGs during nitrate metabolism revealed that many of the genes encoded those involved in protein biosynthetic and metabolic processes, especially ribosome and RNA polymerase biosynthesis, but also other processes including transport and organonitrogen compound metabolism. Analysis of DEGs in the ammonium treatment indicated that cell cycle, oxidoreductase, and certain metabolic activities were reduced. In addition, DEGs participating in the utilization of both nitrate and ammonium were related to L-serine biosynthesis, energy-dependent multidrug efflux pump activity, and glycerol transport. We further showed that the mutants of three differentially expressed transcription factors (*VdMcm1*, *VdHapX*, and *VDAG_08640*) exhibited abnormal phenotypes under nitrate and ammonium treatment compared with the wild type strain. Deletion of *VdMcm1* displayed slower growth when utilizing both nitrogen sources, while deletion of *VdHapX* and *VDAG_08640* only affected nitrate metabolism, inferring that nitrogen assimilation required regulation of bZIP transcription

factor family and participation of cell cycle. Taken together, our findings illustrate the convergent and distinctive regulatory mechanisms between preferred (ammonium) and alternative nitrogen (nitrate) metabolism at the transcriptome level, leading to better understanding of inorganic nitrogen metabolism in *V. dahliae*.

Keywords: *Verticillium dahliae*, nitrogen metabolism, transcriptome, nitrate assimilation, ammonium assimilation

INTRODUCTION

Nitrogen is an indispensable element of biomolecules containing amino acids, nucleotides, and organic cofactors. In all living organisms, nitrogen metabolism requires subtle adjustments to balance metabolite synthesis with uptake from exogenous sources and to ensure survival during conditions of nitrogen starvation (Gouzy et al., 2014). In fungi, ammonium (NH_4^+) and glutamine act as preferred nitrogen sources which are prioritized in their utilization to promote fungal survival and fitness (Ries et al., 2018). In the absence of preferred nitrogen sources, alternative nitrogen sources such as nitrate (NO_3^-) may be utilized causing a switch from nitrogen-associated anabolism to catabolism (Tudzynski, 2014). Nitrogen metabolism plays key roles in the growth and development of plants and microbes, and can affect the fate of an interaction between plants and pathogens (Kusano et al., 2011; Gupta et al., 2013; Thalineau et al., 2016). Acquisition of nitrogen from the host tissues is pivotal to proper infection-related processes because nitrogen supports energy production, provides raw material for growth, and is essential in regulatory or signaling processes in fungal development (Deng et al., 2015).

In plants, nitrogen is transported from the roots to the shoot, mainly as inorganic nitrogen, amino acids, amides, and ureides (Divon et al., 2005). Ammonium and nitrate represent two of the predominant inorganic nitrogen sources that are absorbed by higher plants (Wang M. et al., 2019). The content of these two compounds are diverse in different plant tissues, indicating that the nitrogen sources available for a pathogen in the host plant are dependent on the tissue that is being colonized. Thus it is plausible that for a soilborne pathogen which invades roots and a foliar pathogen, there may be different mechanisms regulating nitrogen metabolism (Snoeijers et al., 2000). For those vascular pathogens that colonize the xylem, the nutrient deprivation for the pathogen may be extreme since the inside of xylem vessels is recognized as a nutritionally poor environment, although it allows efficient transport of water and soluble mineral nutrients from the root to the upper plant parts (Singh et al., 2010). With only minute amounts of solutes (Fisher, 2000), only a small minority of pathogens have adapted to the vascular system of xylem vessels as a major ecological niche in their disease cycles. In the foliar pathogen, *Magnaporthe oryzae*, nitrogen utilization, metabolism of some amino acids, and early infection in the nitrogen-poor confines of the apoplast were studied in depth, revealing that nitrogen sources have an impact on appressorium formation in *M. oryzae* (Froeliger et al., 1996; Wilson et al., 2010; Fernandez and Wilson, 2012; Fernandez et al., 2012; Marroquin-Guzman and Wilson, 2015). Nevertheless, knowledge of nitrogen metabolism of other phytopathogenic fungi is limited.

Vascular wilt diseases caused by pathogenic fungi that proliferate in the water-conducting xylem vessels cause typical symptoms including wilt or flaccidity of the leaves and sometimes death of the plant (Yadeta and Thomma, 2013). For growth, reproduction, and survival, vascular pathogens must readily acclimate to limited-nitrogen environments and efficiently acquire nitrogen resources essential for normal physiological activities in the xylem sap (Divon et al., 2005). Due to the specific niche of vascular pathogens within the xylem, it is difficult to characterize or quantify their adaptations to this environment. Thus, relatively little is known about their metabolism in response to limited nitrogen supplies in the xylem on the molecular and biochemical levels compared to foliar pathogens (Berne and Javornik, 2016). Consequently, it is of theoretical and practical importance to understand the molecular mechanisms of how vascular pathogens adapt to poor nutrition environments and metabolize nitrogen sources from plant hosts so as to design potentially novel control strategies to combat vascular wilt diseases.

Vascular wilt diseases known as Verticillium wilts are caused primarily by the fungal species *V. dahliae*, which infects over 200 plant species worldwide (Klosterman et al., 2009). Verticillium wilts result in yield losses of economically important crops and also affects numerous ornamental shrubs and trees (Klosterman et al., 2009). In China, *V. dahliae* causes high mortality rates of ornamental perennials like smoke trees (*Cotinus coggygria*), the leaves of which provide beautiful scenery at the Fragrant Hills Park in Beijing, China (Wang et al., 2013). Owing to the vascular habitat of the pathogen, no fungicides can provide remedial action once the plant is infected, and control of the pathogen in the field is also difficult since it survives for years in the soil (Klosterman et al., 2009). Though nitrogen metabolism plays indispensable roles in numerous aspects of fungal growth and virulence, there remains a number of questions on nitrogen usage that demand explanation. Relative to the numerous studies on infection structure formation and effectors in *V. dahliae*, the processes of nitrogen acquisition and assimilation in *V. dahliae* have not received adequate attention.

In previous studies, at least two genes in *V. dahliae* were implicated in nitrogen metabolism, including *VdCpc1* which was required for resistance to amino-acid starvation as well as infection and colonization (Timpner et al., 2013). Additionally *VdAtf1* serves as a novel transcription factor (TF) that regulates nitrogen assimilation and further links nitrogen metabolism with virulence (Tang et al., 2020b). Other TFs also take part in nitrogen metabolic processes indirectly (Tang et al., 2020a). However, questions still abound on the influence of nitrogen metabolism on various aspects of the

V. dahliae disease cycle. The purpose of the current study was to understand how *V. dahliae* utilizes diverse inorganic nitrogen sources through transcriptome analysis, especially in a limited nitrogen environment. Our results demonstrate that alternative nitrogen metabolism is less preferred because of the costs associated with the expression of additional gene sets that may not directly influence nitrogen metabolism. In contrast, the preferred nitrogen source stimulated gene expression directly associated with nitrogen metabolism while restraining expression of those genes involved in unrelated cellular and metabolic processes. Results of the data analyses highlight the convergent and distinctive regulatory mechanisms between preferred and alternative nitrogen metabolism, providing an increased understanding of nitrogen metabolism in *V. dahliae*.

MATERIALS AND METHODS

Fungal Strain and Culture Conditions

The wild type *V. dahliae* strain XS11 was isolated from a smoke tree in Fragrant Hills, Beijing (Wang et al., 2013) and was used for transcriptome profile analyses in this study. Mutants including Δ*VdMcm1*, Δ*VdHapX*, Δ*VDAG_08640* were obtained in our previous studies (Xiong et al., 2016; Fang et al., 2017; Wang et al., 2018). Conidial suspensions of all strains were stored long term at −80°C in 30% glycerol. All strains were initially grown on potato dextrose agar (PDA) plates (200 g potato, 20 g glucose, 15 g agar per liter) at room temperature.

For growth tests on TOR and PKA inhibitors, the XS11 strain was cultured for 10 days on complete medium (CM, 50 ml 20 nitrate salts, 1 ml 1000X trace elements, 10 g glucose, 2 g peptone, 1 g yeast extract, 1 g casamino acids, 1 ml vitamin solution per liter) and on defined minimal media containing 1% (w/v) glucose (1 l GMM, 1.52 g KH_2PO_4, 0.52 g KCl, 0.152 g $MgSO_4{\cdot}7H_2O$, 3 μM thiamine HCl, 1.98 g glucose, 15 g agar per liter) with 10 mM of the two sole nitrogen sources nitrate (NO_3^-) and ammonium (NH_4^+) at 25°C, and with 10 μM rapamycin and 40 μM *N*-[2-(*p*-Bromocinnamylamino)ethyl]-5-isoquinolinesulfonamide · 2HCl hydrate (H-89). The colony diameter was measured after 10 dpi (days post inoculation). The inhibited hyphal growth was calculated using difference value of growth diameters of the XS11 strain when cultured on CM and on different nitrogen sources divided by growth diameters of it grown on CM, and those of the XS11 strain treated by rapamycin or H-89 were used the same method of calculation. All of the experiments were repeated three times.

For growth tests on single nitrogen sources, all strains were cultured for 10 days on CM and on GMM with 10 mM of the two sole nitrogen sources nitrate (NO_3^-) and ammonium (NH_4^+) at 25°C. The colony diameter was measured after 10 dpi. The diameters of irregular colonies were measured form multiple different directions and calculated the average value. The inhibited hyphal growth was calculated using difference value of growth diameters of the XS11 strain when cultured on CM and on different nitrogen sources divided by growth diameters of it grown on CM, and those of the mutants were used the same method of calculation. All of the experiments were repeated three times.

RNA Extraction

The 1 ml conidial suspension (10^7 spores/ml) of the XS11 strain was added into 100 ml liquid CM and shaken for 3 days. The fresh hyphae were collected by filtration from the liquid shake cultures using a single-layer of Miracloth (Millipore) and were washed by sterile water three times. Then the vegetative hyphae were switched to liquid GMM with 10 mM of each of the sole nitrogen sources, including nitrate (NO_3^-) or ammonium (NH_4^+) at 25°C for 24 h. The fresh hyphae were collected by filtration from the liquid shake cultures using a single-layer of Miracloth (Millipore) and frozen in liquid nitrogen immediately. All of the fungal samples were ground to a fine powder using a mortar and pestle in liquid nitrogen. Total RNA was extracted by TRIzol reagent (Invitrogen) and purified with a PureLink RNA minikit (Ambion) in accordance with the manufacturers' instructions.

RNA-Seq and Mapping

After checking the quantity and quality of RNA using a NanoPhotometer spectrophotometer (Implen) and an RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies), high quality RNA samples were chosen for RNA-Seq analyses. Three biological replicates under each condition were sequenced by a DNBSEQ-T7 platform (Beijing Genomics Institute), yielding over 45 million reads per sample.

Sequence data filtering was performed using SOAPnuke software (Beijing Genomics Institute). Raw reads were processed using in-house Perl scripts and clean reads were obtained by removing adapter sequences and those reads with low quality and poly-N sequences from the raw data. These high quality data were used to estimate the abundance of each transcript using the *V. dahliae* strain VdLs.17 reference genome from the Broad Institute (Klosterman et al., 2011) and Hierarchical Indexing for Spliced Alignment of Transcripts (HISAT) software (Kim et al., 2015). The reads were aligned to the reference genome (Klosterman et al., 2011) using Burrows-Wheeler transformation (Li and Dewey, 2011) to count the read numbers mapped to each gene. The fragments per kilobase per transcript per million mapped reads (FPKM) values were calculated and used to estimate the effects of sequencing depth and gene length on the mapped read counts. Principal Component Analysis (PCA) analysis and Pearson correlation coefficient matrix were calculated by the princomp function and cor function in R software.

Differential Gene Expression Analysis

DEseq2 (Love et al., 2014) was used to provide statistical analyses for determining differential expression in digital gene expression data. The sets of differentially expressed genes (DEGs) between each pair of libraries were determined by performing Chi-square tests with $P < 0.05$. Significantly expressed genes with a $|\log2(\text{fold-change})| > 1$ and Adjusted *P*-value < 0.05 were selected. The assembled transcripts in the final transcriptome were annotated by mapping them to several public databases.

Gene Ontology (GO) classification of identified DEGs in response to various nitrogen sources was conducted according to functional annotations of DEGs and the GO resource[1]. The DEGs were categorized in terms of biological process, molecular function, and cellular component categories in a GO enrichment analysis, using a GOseq R package (Young et al., 2010), which corrected for gene length bias. GO terms with corrected *P*-values < 0.05 were considered to be significantly enriched within the DEG sets. The corrected *P*-values were adjusted using the Benjamini and Hochberg method (FDR < 0.05) as the threshold for significant differential expression (Benjamini and Yekutieli, 2001). Based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation results, the DEGs were classified into biological pathways. The KEGG database (Hu et al., 2018) was used to identify enriched pathways by a two-tailed Fisher's exact test to test the enrichment of the differentially expressed individual gene products against all those identified. Significant enrichment was determined at $P < 0.05$. These pathways were classified into hierarchical categories according to the KEGG website[2]. Heat maps of expression values were normalized and drawn by TBtools (Chen et al., 2020).

Measurements of Glycerol Content

The mycelia of the XS11 strains were obtained after induction conditions and ground to a fine powder. Subsequently, the powder (0.1 g) of each was transferred to a 2-ml tube containing 1 ml glycerol extraction buffer (Applygen Technologies Inc.). After vortexing for 5 min, the tubes were centrifuged at 5000 *g* for 20 min, 10 μl of supernatant was mixed with 195 μl detection buffer (Applygen Technologies Inc.). After the mixture was incubated at 37°C for 20 min, the glycerol concentration was determined by a spectrophotometer (SpectraMax 190) at 550 nm. The experiment was repeated three times.

Protein Extraction

Fungal samples were ground to a powder in liquid nitrogen and transferred to a 5 ml centrifuge tube. Four volumes of lysis buffer containing 10 mM dithiothreitol, 1% triton-100, and 2 mM ethylenediaminetetraacetic acid (EDTA) (Sigma) were added to the powder, followed by sonication three times on ice using a high intensity ultrasonic processor (Scientz). An equal volume of Tris-phenol was added. Debris were removed following centrifugation at 5000 × *g* at 4°C for 10 min. The protein precipitate was washed with methanol and acetone respectively and was redissolved in 8 M urea. The protein concentration was determined with a BCA kit (Protein Assay Kit) (Beyotime) according to the manufacturer's instructions.

Statistical Analysis

Data were expressed as mean value ± standard error of the mean. Statistical analyses were performed by ANOVA and Student's *t*-test (SPSS 16.0). The *P*-value < 0.05 and <0.01 was considered statistically significant in this study.

[1]http://geneontology.org/

[2]https://www.kegg.jp/kegg/kegg2.html

Availability of Data and Material

Raw sequences were deposited in the National Center for Biotechnology Information (NCBI) and can be accessed in the Short Read Archive (SRA) database[3] under accession SRX11048827-SRX11048835 for XS11, XS11-NO_3^-, and XS11-NH_4^+ respectively.

RESULTS

RNA-Seq Profiles of *Verticillium dahliae* in Response to Nitrate and Ammonium Treatments

Xylem vessels of plants allow efficient transport of some soluble nitrogen sources such as ammonium and nitrate from the root to the upper plant parts, which can be used by *V. dahliae* as nutrients. Nevertheless, it is a nitrogen-deficient environment, and *V. dahliae* must acclimate rapidly for survival. To explore transcriptomes of *V. dahliae* on various inorganic nitrogen sources, three biological replicates of the XS11 strain were harvested for transcriptome sequencing using RNA-seq for each condition (non-treated, 10 mM nitrate, and 10 mM ammonium). These nine RNA-seq datasets were sequenced using DNBSEQ-T7, and approximately 47 million reads were obtained for each of the libraries. By removing adaptor sequences and undesirable reads (ambiguous, low quality, and duplicated sequence reads), clean reads were generated from the nine libraries with Q30 > 85%, suggesting high quality for the present sequencing results (**Table 1**). The total number of clean reads per library ranged from 42.09 to 44.78 million, and clean bases ranged from 6.31 to 6.72 Gb. Clean reads from all libraries were aligned to annotated genes in the *V. dahliae* VdLs.17 genome (Klosterman et al., 2011), and the fragments per kilobase per transcript per million mapped reads (FPKM) value was calculated for each gene.

The relationships among the nine samples were illustrated using a principal component analysis (PCA) performed with log2-normalized FPKM of all expressed genes. The samples from the same treatments were coordinated together, as expected, and the samples from the different treatments were clearly differentiated (**Figure 1A**). The Pearson's correlation coefficients of paired samples were calculated using the FPKM values of all expressed genes for examination of the correlations between gene expression patterns among the nine libraries. All of the Pearson's correlation coefficients were > 0.6. We also found the strong correlation among the samples of the same treatment (R^2 = 0.996–1, $P < 0.001$), indicating a high level of reproducibility of the RNA-seq data (**Figure 1B**). Meanwhile, the samples obtained from different processing conditions yielded much lower correlation coefficient values. These results illustrated that samples in the same treatment groupings had consistency in their transcriptional changes and repeatability among the RNA-seq profiles.

[3]http://trace.ncbi.nlm.nih.gov/Traces/sra/

TABLE 1 | RNA sequencing statistics.

Sample	Total raw reads (M)[a]	Total clean reads (M)[b]	Clean reads Q30 (%)[c]	Total mapping (%)[d]	Uniquely mapping (%)[e]
XS11_1	49.08	44.78	85.28	90.91	62.83
XS11_2	47.33	43.51	85.91	90.97	63.23
XS11_3	47.33	43.18	85.19	91.06	62.48
XS11-NO3_1	47.33	43.84	88.12	91.75	68.26
XS11-NO3_2	47.33	43.89	87.94	91.54	68.03
XS11-NO3_3	45.57	42.09	87.51	91.49	66.78
XS11-NH4_1	47.33	43.45	86.4	91.15	67.39
XS11-NH4_2	47.33	43.42	86.44	91.06	66.66
XS11-NH4_3	47.33	43.49	86.42	91.22	66.8

[a] *Number of reads generated from the BGISEQ-500.*
[b] *Raw data with low-quality reads, reads containing 3' adaptors, and reads containing two or more N bases removed.*
[c] *Percentage of bases with a mass value > 30.*
[d] *Number of reads mapped to the reference genome within 2-bp mismatch.*
[e] *Uniquely mapped = the number of reads mapped to the reference genome with unique sequence location.*

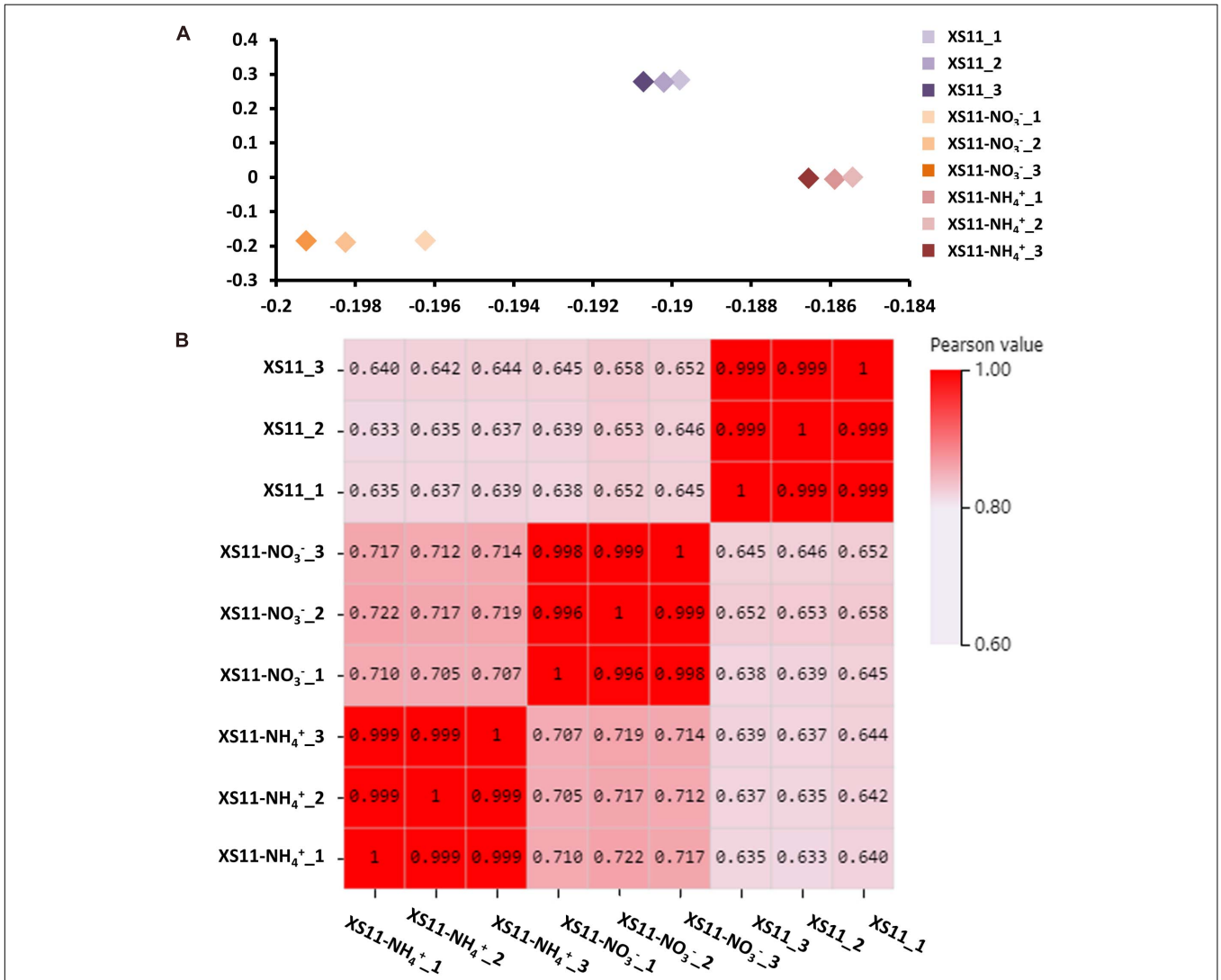


FIGURE 1 | Repeatability of RNA-seq expression profiles in biological replicates in *Verticillium dahliae* in response to two different nitrogen sources. **(A)** Principle component analysis (PCA) of expression data about samples with or without treatments including 10 mM nitrate (NO_3^-) or 10 mM ammonium (NH_4^+) with three biological replicates. **(B)** Pearson's correlation values ($P < 0.01$) non-clustering heatmap of nine samples displaying the correlation coefficient (R^2).

Identification of Differentially Expressed Genes (DEGs) in Response to Nitrate or Ammonium

Analyses of the DEGs collected may provide increased understanding into the regulatory changes in *V. dahliae* using different inorganic nitrogen sources. For this reason, we first identified all DEGs, and examined the major differences within these numbers between the nitrate or ammonium treatments. In total, expression of 8585 genes was altered between the two treatments. From a global overview, 7214 genes were affected under nitrate treatment including 3633 that were upregulated and 3581 that were downregulated in their expression (**Figure 2A**). Additionally, 6962 genes were influenced when ammonium served as the nitrogen source, with 3440 genes that were upregulated and 3522 that were downregulated. This finding suggested that *V. dahliae* responded by increased differential gene expression under nitrate treatment as compared to ammonium treatment.

Genes with insignificant differential expression values were filtered out, and DEGs ($|\log2(\text{fold-change})| > 1$ and Adjusted *P*-value < 0.05) were kept for further analyses. Based on the volcano plot analysis, 3244 genes were significantly differentially expressed in the XS11-NO_3^- versus XS11 comparison, which represented nearly half of the affected genes in the nitrate treatment group (**Figure 2B**). Among these, 1840 DEGs were increased in their expression and 1404 were markedly reduced in their expression. While the XS11-NH_4^+ versus XS11 comparison revealed 922 upregulated and 1606 downregulated DEGs (**Figure 2C**). The up- and down-regulated DEGs were of about the same number in response to nitrate treatment, while the vast majority of DEGs were as downregulated in the ammonium treated group.

Gene Ontology (GO) Analyses of DEGs

To understand the function of the DEGs uncovered in the nitrate and ammonium treated groups, the DEGs were distributed into three GO categories including biological process (BP), cellular component (CC), and molecular function (MF). Under both treatments, GO terms "catalytic activity," "membrane," and "metabolic process" accounted for the greatest number in MF, CC, and BP, respectively (**Figure 3**). DEGs of the XS11 strain treated by both nitrogen sources exhibited similar GO classifications and ratios. However, "molecular carrier activity" containing VDAG_04520 (encoding Cytochrome C Oxidase Copper Chaperone, COX17) VDAG_10000 (encoding exportin-1, XPO1) only appeared in the nitrate treatment group.

For further study, all DEGs were analyzed by GO enrichment. In the BP category, DEGs in response to the nitrate treatment were significantly enriched in 29 GO terms (Corrected *P*-value < 0.01), and the top five terms significantly enriched were macromolecule metabolic process (GO: 0043170); cellular nitrogen compound metabolic process (GO: 0034641); cellular macromolecule metabolic process (GO: 0044260); gene expression (GO: 0010467); and cellular macromolecule biosynthetic process (GO: 0034645) (**Supplementary Table 1**). Likewise, DEGs in response to ammonium treatment were significantly enriched in 59 GO terms (Corrected *P* value < 0.01), and the top five terms significantly enriched were organic cyclic compound metabolic process (GO: 1901360); cellular aromatic compound metabolic process (GO: 0006725); heterocycle metabolic process (GO: 0046483); nucleobase-containing compound metabolic process (GO: 0006139); and nucleic acid metabolic process (GO: 0090304). DEGs were gathered in different locations under different treatments in the CC category, non-membrane-bounded organelle, ribonucleoprotein complex, and ribosome for nitrate treatment whereas membrane and nucleus for ammonium treatment. GO enrichment of DEGs in the two treatments also showed similar and distinct differences in the MP category. Oxidoreductase activity (GO: 0016491) was the most significantly enriched term in response to both treatments, and DEGs of nitrate-treated XS11 were enriched in transporter activity (GO: 0005215) and transmembrane transporter activity (GO: 0022857) while DEGs of ammonium-treated XS11 were enriched in flavin adenine dinucleotide binding (GO: 0050660) (**Supplementary Table 1**).

Kyoto Encyclopedia of Genes and Genomes (KEGG) Analyses of DEGs

The distribution of KEGG terms and their ratios in the six categories for each treatment are shown in **Figure 4**. A total of 3244 (nitrate) and 2528 (ammonium) genes were mapped to 22 and 23 KEGG pathways, respectively, and most of them were overlapping. The analysis showed that in the global and overview maps, carbohydrate metabolism, signal transduction, transport and catabolism, and amino acid metabolism were among the top five terms in all pathway categories under both treatment groups. "Drug resistance: Antimicrobial" pathway containing VDAG_05701 only appeared as enriched in response to ammonium treatment.

Furthermore, KEGG enrichment was applied to identify pathways associating with the DEGs with diverse nitrogen utilization patterns (**Supplementary Table 2**). The KEGG enrichment revealed similar results to those of the GO enrichment, indicating that more genes related to ribosomes and transporters are upregulated in the nitrate treatment. Expression of genes related to terms ribosome (ko03010) and RNA polymerase (ko03020) were all highly upregulated (**Figure 5**), while ABC transporters (ko02010) were downregulated in response to the nitrate treatment. Enzymes involved in lipid metabolism, carbohydrate metabolism, and amino acid metabolism were abundant when assimilating and metabolizing ammonium, accounting for DEGs involved in these significantly enriched metabolic pathways.

Considering the link between TOR and PKA pathways and ribosomal DNA transcription (Lippman and Broach, 2009), rapamycin was used to inhibit the TOR pathway and *N*-[2-(*p*-Bromocinnamylamino)ethyl]-5-isoquinolinesulfonamide · 2HCl hydrate (H-89) was used to inhibit Protein kinase A (PKA) when the XS11 strain metabolized nitrate and ammonium to examine the connection between ribosomal DNA transcription and nitrogen metabolism. Rapamycin led to higher inhibition rates under both treatments, while H-89 activated growth of the strain when it utilized nitrate (**Figure 6**). The expression of the

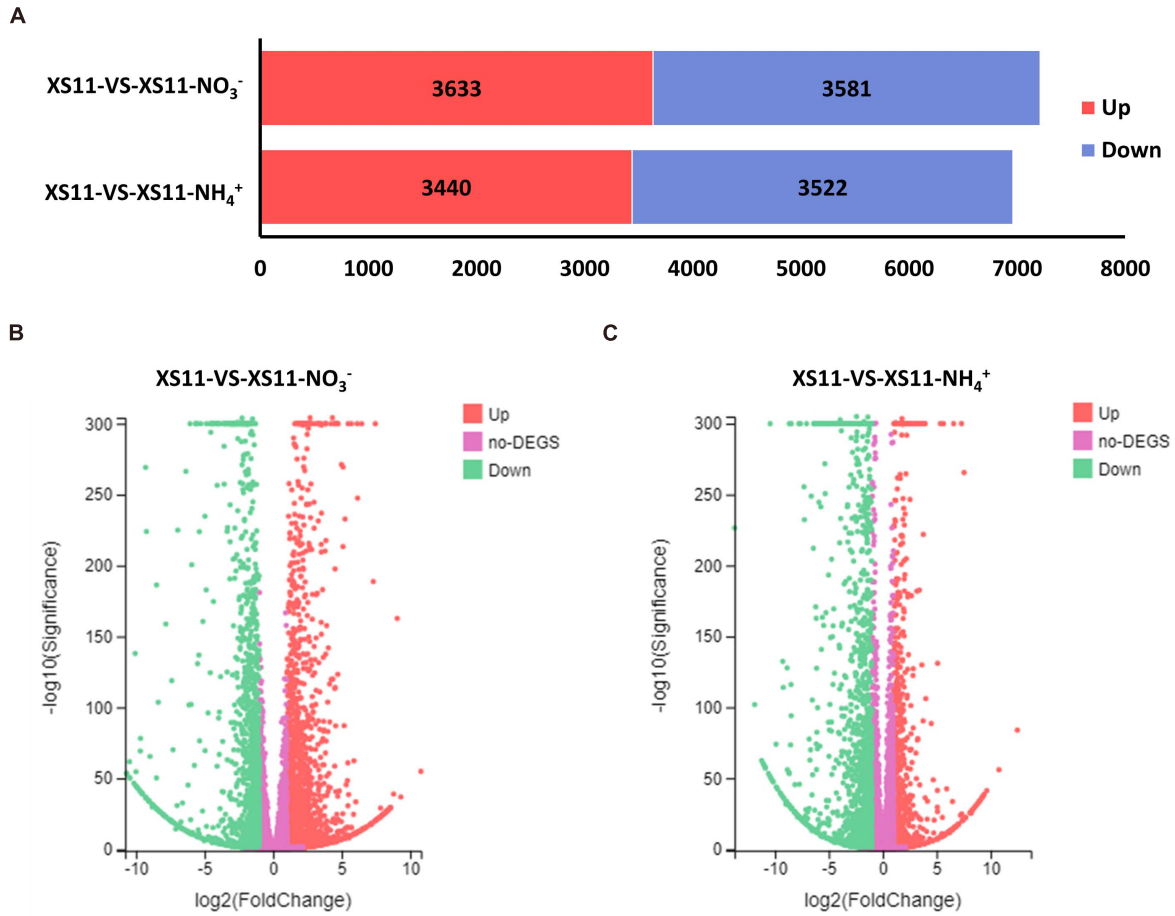


FIGURE 2 | The number of differentially expressed genes (DEGs) in *Verticillium dahliae* in response to different nitrogen treatments. **(A)** Bar chart showing a different number of DEGs in the XS11 strain under different treatments in this study. The treatments included 10 mM nitrate (NO_3^-) or 10 mM ammonium (NH_4^+) amendments, or control. The red color represents gene upregulation while blue represents downregulation. **(B)** and **(C)**. Volcano plot showed the distribution of DEGs at | log2(fold-change) | > 1 and FDR $P < 0.05$ in the *V. dahliae* XS11 strain in response to 10 mM nitrate **(B)** or 10 mM ammonium **(C)** versus the control. The red point in the plot represents the upregulated DEGs with statistical significance, the green represents downregulated DEGs with statistical significance, and purple indicates DEGs without statistical significance.

key genes involved in TOR and PKA pathways were not changed significantly, inferring nitrogen metabolism impacted ribosomal DNA transcription through protein function rather than at the transcription level. Higher inhibition owing to restrained ribosomal DNA transcription may restrict nitrogen utilization of *V. dahliae*.

Convergent and Distinctive Regulatory Mechanisms for Nitrate and Ammonium Metabolism

To reveal the similarities and differences in gene expression during nitrate or ammonium utilization in *V. dahliae*, we applied Venn diagrams to profile the DEG distribution between XS11-NO_3^- versus XS11 and XS11-NH_4^+ versus XS11. In the upregulated DEGs, 1457 DEGs were only significantly expressed in XS11-NO_3^- versus XS11 comparison, which was nearly three times more than those only in XS11-NH_4^+ versus XS11, and 383 DEGs were highly upregulated in both groups (**Figure 7A**). However, the number of DEGs downregulated in XS11-NO_3^- versus XS11 comparison was slightly lower than those downregulated in the XS11-NH_4^+ versus XS11 comparison (**Figure 7B**). The number of DEGs downregulated in both groups was nearly twice than that upregulated in both groups. VDAG_08640 was the top ten genes activated in both groups, and VDAG_05022 (*VdHapX*) was downregulated in response to both treatments. The mutant strains exhibited inhibited growth in nitrate utilization after deletion of the two genes, though they were similar to the XS11 strain when using ammonium (**Figures 7C,D**). VDAG_08640 and *VdHapX* might participate in nitrate metabolism other than ammonium utilization.

GO and KEGG enrichments were conducted for analyses of up and downregulated DEGs overlapped in the XS11-NO_3^- versus XS11 and the XS11-NH_4^+ versus XS11 comparisons, respectively. There were no significant enrichments in upregulated DEGs and KEGG enrichment while DEGs which

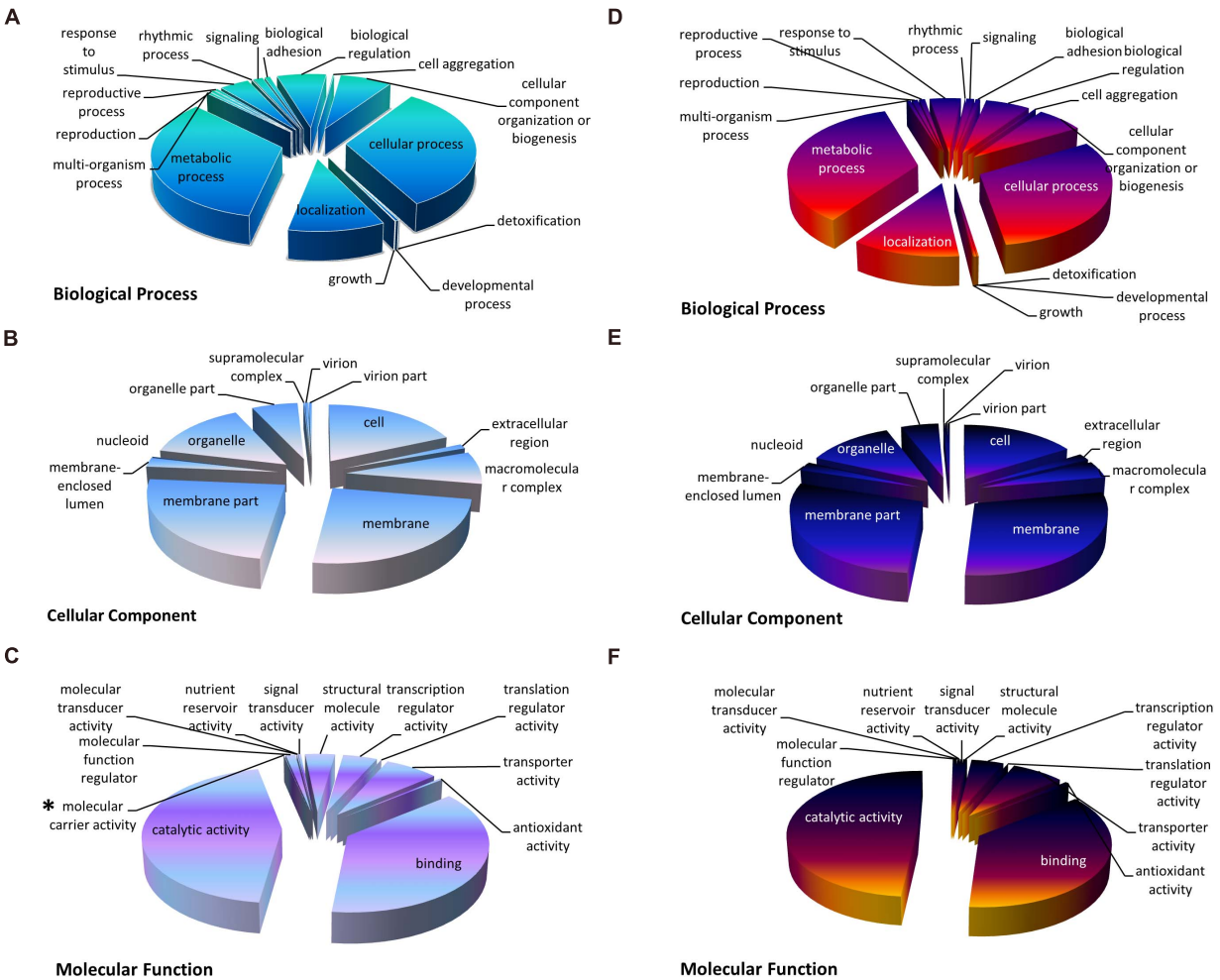


FIGURE 3 | The Gene Ontology (GO) classification of DEGs under different nitrogen treatments in *Verticillium dahliae*. GO classification of identified DEGs based on their functional annotations in response to various nitrogen sources was conducted according to the GO resource (http://geneontology.org/). The treatments included 10 mM nitrate (NO_3^-) or 10 mM ammonium (NH_4^+) amendments, or control. **(A–C)** Indicates biological processes, cellular components, and molecular function under nitrate treatment, while **(D–F)** indicate biological processes, cellular components, and molecular function under ammonium treatment, respectively. *Indicates that the term appeared only once.

were downregulated in both groups were collected into 15 GO terms and more than one-third of these DEGs corresponded to the location term of membrane part. These also included terms of oxidoreductase activity (GO: 0016491); small molecule metabolic process (GO: 0044281); carbohydrate metabolic process (GO: 0005975); carboxylic acid metabolic process (GO: 0019752); oxoacid metabolic process (GO: 0043436); and alpha-amino acid metabolic process (GO: 1901605); transporter activity (GO: 0005215); transmembrane transporter activity (GO: 0022857); active transmembrane transporter activity (GO: 0022804); and copper ion transmembrane transporter activity (GO: 0005375) (**Figure 8A**). KEGG enrichment analysis of downregulated DEGs exhibited no significantly enriched pathway.

In the XS11-NO_3^- versus XS11 group, upregulated and downregulated DEGs were significantly enriched in 34 and 16 GO terms, respectively. DEGs upregulated in the nitrate utilization group included those associated with terms non-membrane-bounded organelle and focused on nitrogen metabolism, including cellular nitrogen compound biosynthetic process (GO: 0044271), organonitrogen compound biosynthetic process (GO: 1901566); organonitrogen compound metabolic process (GO: 1901564); relevant processes of gene transcription and translation including gene expression (GO: 0010467); translation (GO: 0006412); ribosome biosynthesis such as ribonucleoprotein complex (GO: 1990904); ribosome (GO: 0005840); structural constituent of ribosome (GO: 0003735); peptide and protein biosynthetic and metabolic processes such as peptide metabolic process (GO: 0006518); peptide biosynthetic process (GO: 0043043); cellular protein metabolic process (GO: 0044267); protein metabolic process (GO: 0019538); other cellular process containing amide biosynthetic process (GO: 0043604); and cellular amide metabolic process (GO: 0043603) (**Figure 8B**). Additionally, KEGG enrichment analysis indicated one strikingly enriched pathway of ribosome (ko03010),

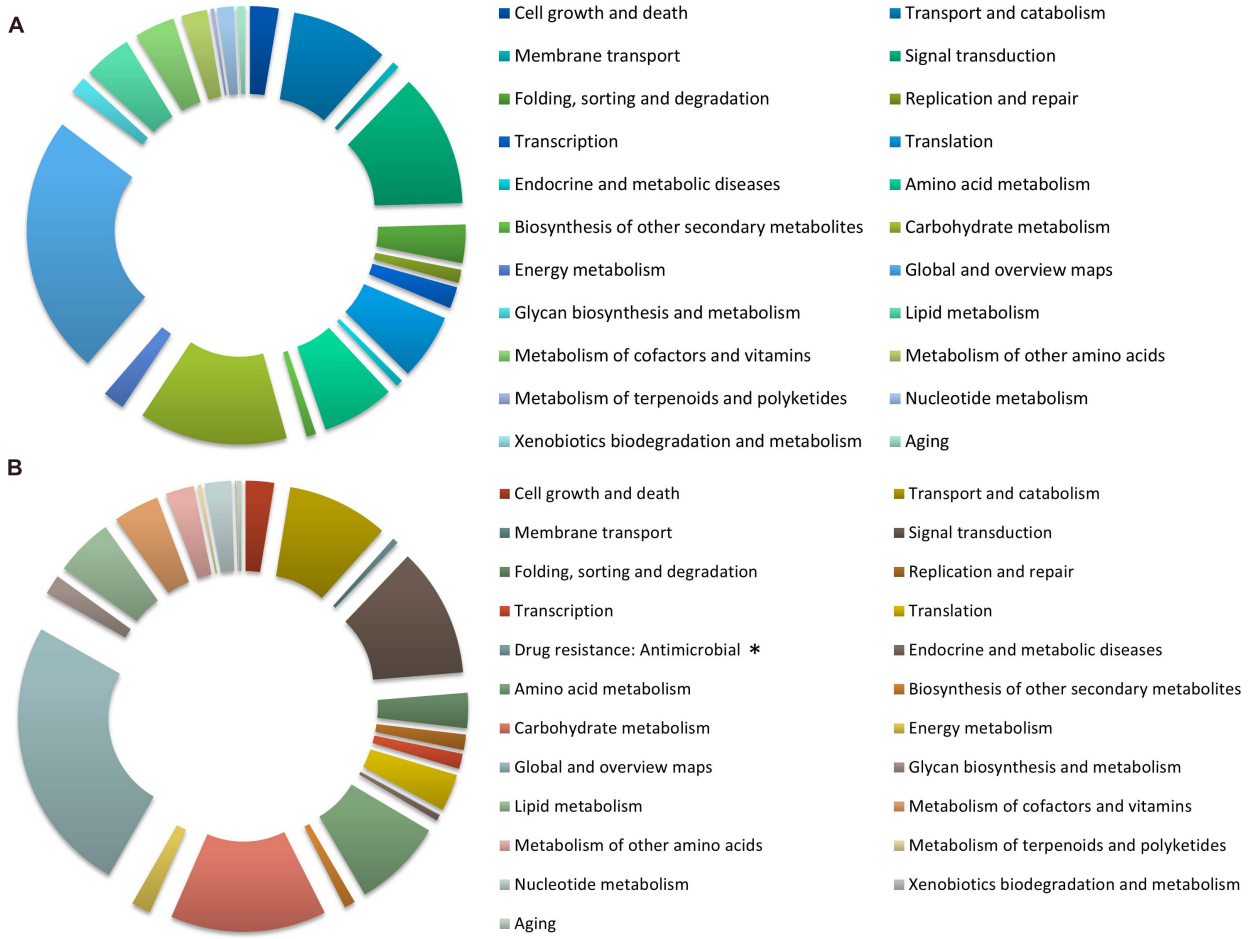


FIGURE 4 | The Kyoto Encyclopedia of Genes and Genomes (KEGG) classification of DEGs under different nitrogen treatments in *Verticillium dahliae*. The pie charts exhibit KEGG classification of genes differentially expressed in *V. dahliae* strain XS11 in response to various nitrogen sources, including 10 mM nitrate (NO_3^-) or 10 mM ammonium (NH_4^+). **(A)** Represents XS11-NO_3^- versus XS11 and **(B)** represents XS11-NH_4^+ versus XS11. * Indicates that the pathway appeared only once. These pathways were classified into hierarchical categories according to the KEGG website (https://www.kegg.jp/kegg/kegg2.html).

which was consistent with the GO enrichment of upregulated DEGs. DEGs downregulated in this group were related to functions in binding ions and anions; adenyl nucleotide; adenyl ribonucleotide; ATP; GTPase; signal transduction such as Ras protein signal and small GTPase mediated signal transduction; active transmembrane transporter activity; and cell communication (**Figure 8C**). But KEGG enrichment analysis showed no significantly enriched pathway. *V. dahliae* activated numerous genes participating in protein biosynthesis when utilizing nitrate, and protein accumulated more in nitrate-treated XS11 vs. the ammonium-treated strain XS11 (**Figure 8D**).

In the XS11-NH_4^+ versus XS11 group, upregulated and downregulated DEGs were significantly enriched in 13 and 5 GO terms, respectively. Among the upregulated DEGs, we identified a substantial number of those associated with RNA processes containing RNA biosynthetic process (GO: 0032774); RNA metabolic process (GO: 0016070); ncRNA processing (GO: 0034470); rRNA processing (GO: 0006364); and rRNA metabolic process (GO: 0016072); transcription such as transcription, DNA-templated (GO: 0006351); and nucleic acid-templated transcription (GO: 0097659); nuclear and ribosome including nucleolus (GO: 0005730); nucleic acid metabolic process (GO: 0090304); ribonucleoprotein complex biogenesis (GO: 0022613); preribosome (GO: 0030684); and ribosome biogenesis (GO: 0042254) (**Figure 9A**). Most of these were located in the compartment of the nucleolus. KEGG enrichment analysis showed no significantly enriched pathway in upregulated DEGs. Moreover, downregulated DEGs mostly localization as integral component of membrane and were enriched in oxidoreductase and peptidase activities (**Figure 9B**). Considering activation of genes involved in ribosome biogenesis and depression of those related to peptidase activity, we measured protein content when using ammonium, showing obvious protein accumulation (**Figure 8D**). KEGG enrichment analysis displayed pathways related to cell cycle (ko04111) and metabolism relating to lipid metabolism (ko00565); glycerophospholipid metabolism (ko00564); pyrimidine metabolism (ko00240); and glycine, serine, and threonine metabolism (ko00260). *VdMcm1*

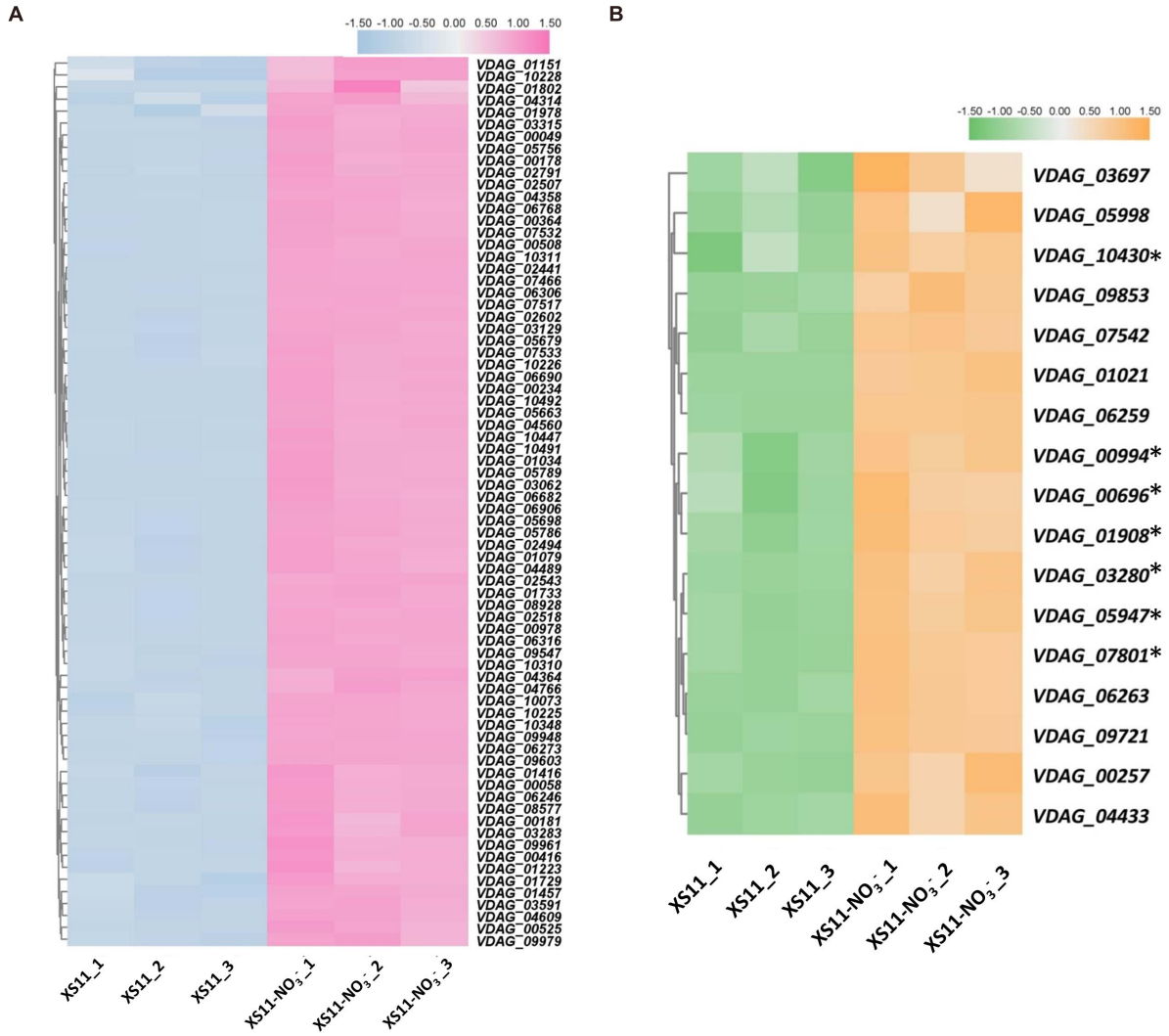


FIGURE 5 | Heat maps of differentially expressed genes related to ribosome and RNA polymerase processes in response to nitrate as the sole nitrogen source in *Verticillium dahliae*. **(A)** Heat map showing the relative expression of genes related to ribosome in the *V. dahliae* XS11 strain using nitrate (NO_3^-) and in the control. TBtools was used to normalized data and generate heat map. The pink color indicates expression was relatively higher, while blue color indicates expression was relatively lower. White indicates that the gene expression between conditions analyzed was unchanged. **(B)** Heatmap of comparing expression of genes involved in RNA polymerase in nitrate metabolism. TBtools was used to normalized data and generate heat map. The orange color in the heat map indicates expression was relatively higher, while the green color in the heat map indicates expression was relatively lower. The white color indicates that the gene expression between conditions analyzed was unchanged. *Means the key genes in RNA polymerase.

(*VDAG_01770*) is a conserved transcription factor regulating the cell cycle, so the Δ*VdMcm1* strain was tested here to ascertain its role in nitrogen utilization. Its inhibition rate was higher than the XS11 strain, which meaning *VdMcm1* takes part in regulation of nitrogen metabolism (**Figures 7C,D**). Genes involved in glycerophospholipid metabolism were depressed during ammonium utilization, which might cause perturbations in the glycerol content in hyphae. However, we found that ammonium treatment did not change glycerol content, while nitrate treatment caused its increase (**Figure 9C**). Our data showed that expression of VDAG_05602 (glycerol kinase, GK) was depressed in response to both treatments and the expression of VDAG_00445 (triacylglycerol lipase, TGL) was upregulated only when treated with nitrate, which might explain the differences in glycerol content in different nitrogen sources.

DISCUSSION

Nitrogen availability plays an essential role in growth, basic metabolic processes, and even plant-pathogen interactions for fungal pathogens (Kusano et al., 2011; Gupta et al., 2013; Thalineau et al., 2016). For acquisition and metabolism of different nitrogen sources, fungal pathogens have developed sophisticated mechanisms. There is a prioritization to use preferred nitrogen sources while repressing the utilization of

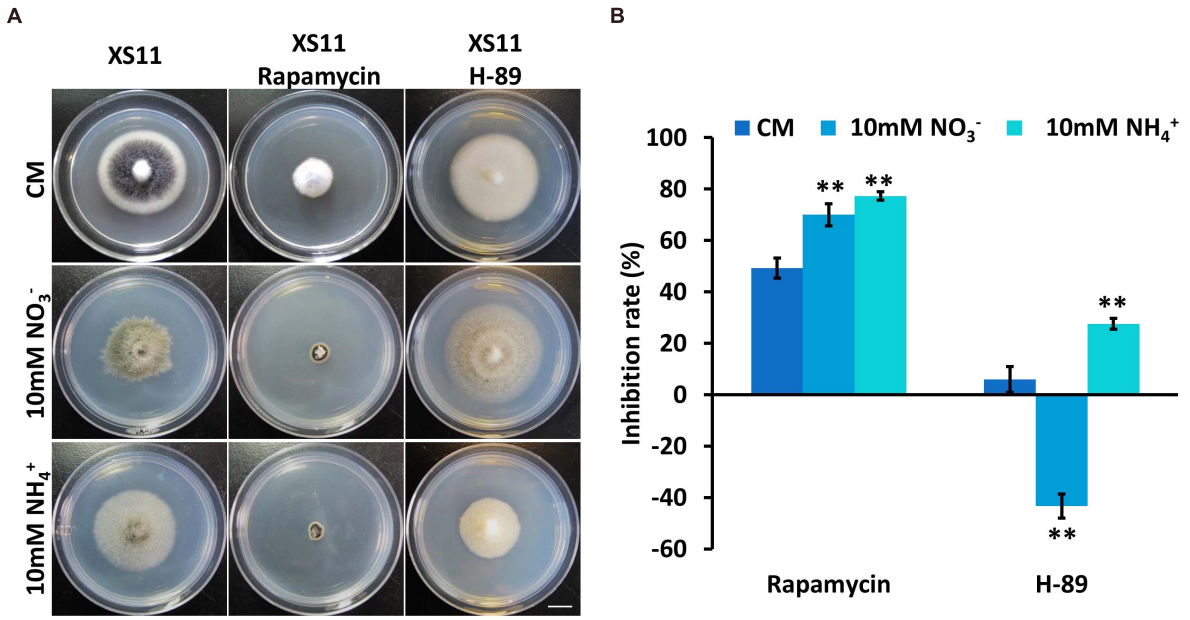


FIGURE 6 | Responses of *Verticillium dahliae* to rapamycin and H-89 when utilizing two different nitrogen sources. **(A)** The XS11 strain was grown for 10 days on CM and on defined minimal media containing 1% (w/v) glucose (GMM) with 10 mM of the indicated sole nitrogen source including NO_3^- and NH_4^+, and with 10 μM rapamycin and 40 μM *N*-[2-(*p*-Bromocinnamylamino)ethyl]-5-isoquinolinesulfonamide · 2HCl hydrate (H-89) at 25°C. Photographs were taken at 10 days post inoculation (dpi) of the media. Scale = 1.05 cm. **(B)** The chart illustrates the inhibition rates of the above colonies at 10 dpi. Error bars represent the standard deviation based on three independent replicates. Asterisks indicate significant differences (**$P < 0.01$).

alternative nitrogen sources, but utilizing alternative nitrogen sources causes switching from nitrogen-associated anabolism to catabolism (Tudzynski, 2014). Transcriptome analyses in this study showed clear similarities and differences between gene expression patterns in response to preferred or alternative nitrogen source utilization in *V. dahliae*. The higher numbers of different DEGs in the nitrate (NO_3^-) treatment versus the ammonium (NH_4^+) treatment suggested that alternative nitrogen sources require additional energy in metabolism, while preferred nitrogen sources such as ammonium are assimilated and metabolized in association with repression of certain cellular processes.

Variation of Gene Expression in *V. dahliae* in Response to Preferred or Alternative Nitrogen Sources

Our previous study revealed that *V. dahliae* grew more slowly when nitrate was supplied as the sole nitrogen source vs. ammonium (Tang et al., 2020b). Herein, these transcriptome data provide evidence that may explain the differences in growth in response to either nitrogen or ammonium as the sole nitrogen source. As a whole, the two nitrogen metabolism affects the expression of genes with different functions and at different cellular sites or compartments. Nitrate utilization regulates genes located at ribosomes and non-membrane-bounded organelles and takes part in macromolecule and cellular nitrogen compound metabolic processes, gene expression, and transporter activity. However, ammonium utilization changes expression of genes located in membrane and nucleus and participates in some organic compound and nucleic acid metabolic process. During the utilization of nitrate shifts in metabolic processes are evident, and some these may be undesirable, but the choices of different compounds may be limited for survival. Additionally, the expression of genes involved in cellular organonitrogen compound biosynthetic and metabolic processes rose in response to greater extent during nitrate metabolism treatment rather than ammonium metabolism, suggesting that these genes are activated in response to alternative nitrogen metabolism.

Nitrate and ammonium metabolic processes employ different regulatory mechanisms despite their similar cellular processes. For example, among all upregulated DEGs in response to nitrate and ammonium the majority were enriched in transcriptional and translational processes, indicating the requirement by the pathogen to increase gene expression and protein translation to utilize nitrogen sources, which was evident by the accumulated protein content we observed in the nitrogen-treatment. The predicted locations of these products differed by organelle part: those in the nitrate treatment were localized in the ribosomal subunit and mitochondrial intermembrane space while those upregulated DEGs from the ammonium were mostly localized in nucleolus. These upregulated sets also have a bias toward distinct processes. The nitrate treatment correlated with increased numbers of DEGs associated with ribosome biosynthesis as well as protein biosynthetic and metabolic processes, while ammonium treatment resulted in upregulated DEGs that favored RNA and transcription related processes.

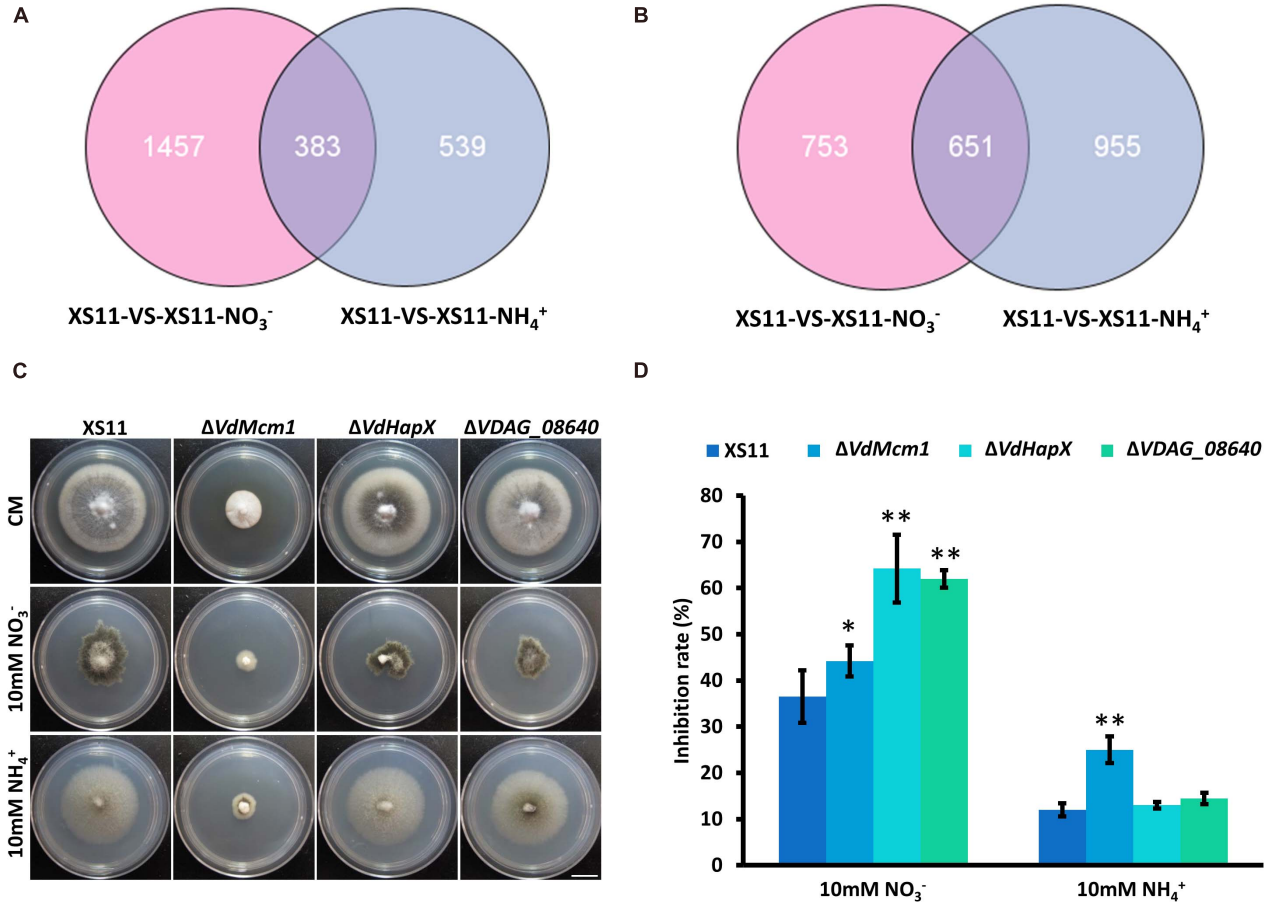


FIGURE 7 | Upregulated and downregulated differentially expressed genes (DEGs) in response to two different nitrogen sources in *Verticillium dahliae*. **(A)** The Venn diagram shows the number of DEGs upregulated in the treatments of strain XS11 of *V. dahliae*. The treatments included 10 mM nitrate ($NO3^-$) or 10 mM ammonium ($NH4^+$) amendments, and control. The sum of numbers in each circle represents the total number of DEGs between the two samples; overlapping parts of the circle represent commonly expressed DEGs between the two treatments. Scale = 1.33 cm. **(B)** Venn diagram showing the number of DEGs downregulated in treatments. The sum of numbers in each circle represents the total number of DEGs between the two samples; overlapping parts of the circle represent commonly expressed DEGs between the two treatments. **(C)** All strains were grown for 10 days on CM and GMM with 10 mM of the indicated sole nitrogen source including NO_3^- and NH_4^+ at 25°C. Photographs were taken at 10 dpi of the media. **(D)** The chart illustrates the inhibition rates of the above colonies at 10 dpi. Error bars represent the standard deviation based on three independent replicates. Asterisks indicate significant differences (**$P < 0.01$; *$P < 0.05$).

Ribosomes are primary consumers of cell resources and energy, thus accurate control of their biogenesis is fundamental for the maintenance of cellular growth. The transcription of ribosomal DNA is a key step in ribosome biosynthesis and a major biosynthetic and energy consuming process, which is tightly coupled to the availability of growth factors, nutrients and energy (Kusnadi et al., 2015). In a previous study, inhibition of mTORC1 with rapamycin led to reduced ribosomal DNA transcription, and it might related to amino acid availability (James and Zomerdijk, 2004; Hannan et al., 2013). Ribosome biosynthesis also depends on the balance between the economics of protein production and cell growth (Slavov et al., 2015). Pathogens may actively downregulate the expression of ribosomal genes to inhibit the translation system for efficient expression and translation of genes related to stress response under stress conditions (Spriggs et al., 2010). When confronted with nitrogen limitation, 17 selective ribosomal proteins were upregulated for activation of reserve translational capacity, revealing large reserves in metabolic and translational capacities of yeast (Yu et al., 2020). Rapamycin reduced hyphae growth when utilizing the two nitrogen sources herein, suggesting that the metabolism of the alternative nitrogen source requires normal ribosomal DNA transcription regulated by TOR, while the same process during ammonium utilization may be regulated by amino acid metabolism. PKA can also activate genes for ribosomal biogenesis (Lippman and Broach, 2009), however, inhibition of PKA led to higher growth rate during utilizing nitrate. DEGs related to signal transduction such as Ras protein signal and small GTPase mediated signal transduction were downregulated in nitrate utilization, while expression of genes encoding PKA did not changed significantly. Inhibition of PKA caused a corollary change in signal transduction related genes that matched those associated with PKA activity, potentially facilitating growth. It may be inferred that there was not

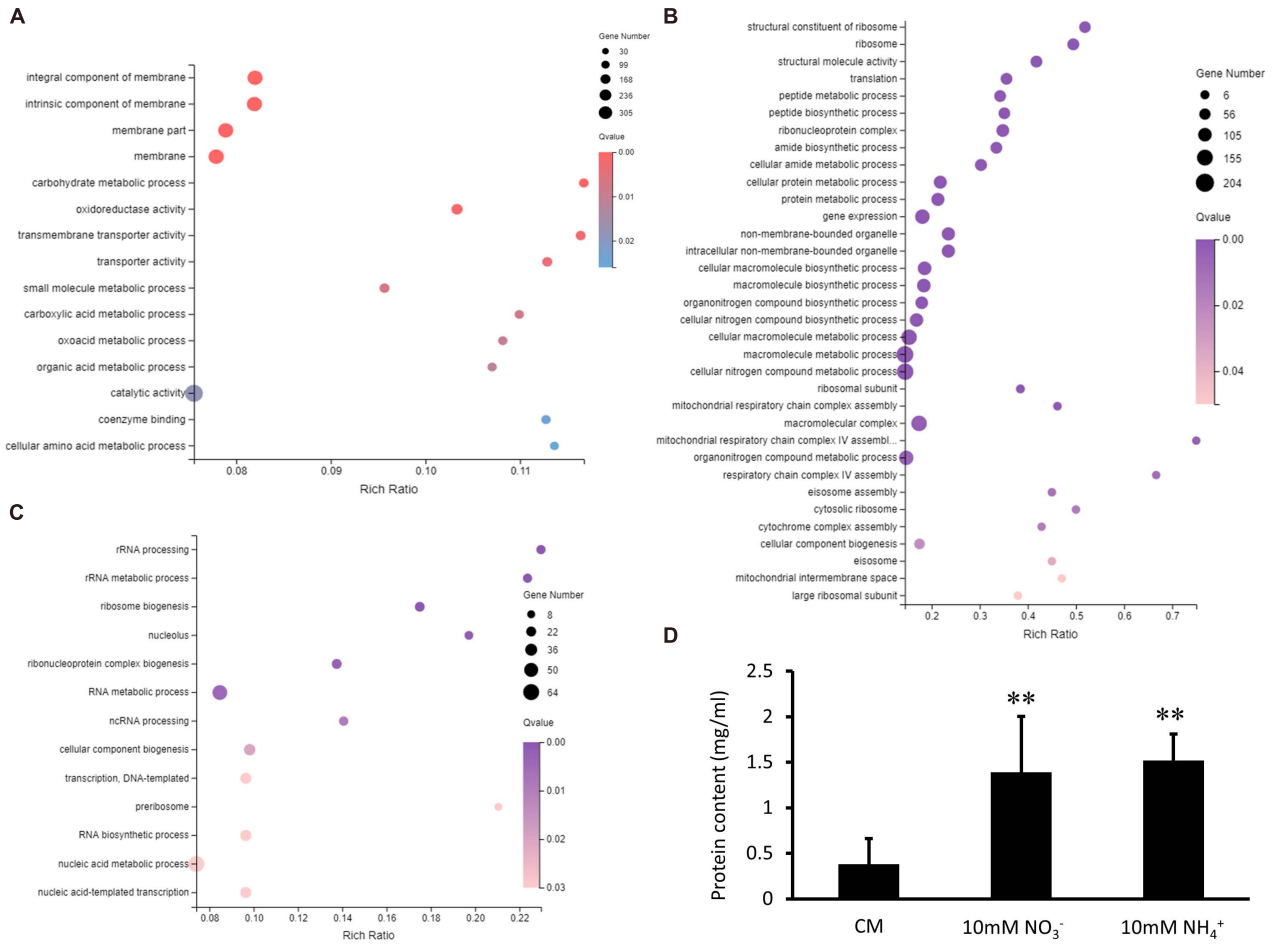


FIGURE 8 | Enrichment analyses of differentially expressed genes (DEGs) in *Verticillium dahliae* in response to different nitrate sources. **(A)** A bubble chart showed DEGs depressed in both groups were enriched in 15 gene ontology (GO) terms. The size of the bubbles is proportional to the total number of DEGs. The *X*-axis represents the enrichment ratio and the *Y*-axis represents GO term. **(B)** A bubble chart showing DEGs whose expression was increased in nitrate utilization and enriched in 34 gene ontology (GO) terms by using a GOseq R package mentioned in section "Materials and Methods." The size of the bubbles is proportional to the total number of DEGs. The *X*-axis indicates the enrichment ratio, and the *Y*-axis indicates the GO terms. **(C)** A bubble chart showing DEGs downregulated in nitrate utilization and enriched in 16 GO terms. The size of the bubbles is proportional to the total number of DEGs. The *X*-axis indicates enrichment ratio, and the *Y*-axis indicates GO terms. **(D)** Bar chart showing the protein content of the XS11 strain treated with 10 mM $NO3^{-}$ or 10 mM $NH4^{+}$ for 24 h. Error bars represent the standard deviation based on three independent replicates. Asterisks indicate significant differences (**$P < 0.01$; *$P < 0.05$).

enough ATP for PKA when nitrate was utilized, causing inhibition of growth.

Over-accumulation of glycerol in the cell may lead to various morphological defects and ultimately growth inhibition, though its normal accumulation is a cellular adaptation to high osmotic stress and salt stress (Randhawa et al., 2019). Fungi must maintain glycerol levels within limits suitable for growth, and tightly control its production and suppression (Kayingo et al., 2001). Glycerol uptake is also an important physiological mechanisms associated with glycerol utilization and osmoregulation (Ji et al., 2018). Stl1, a sugar transporter contributing to glycerol import, was first identified in *S. cerevisiae* to be involved in active glycerol uptake (Ferreira et al., 2005). It is also involved in osmotic response in *S. cerevisiae* and *C. albicans* (Kayingo et al., 2009; Bai et al., 2015). Stl1 was expressed in a complex manner by carbon source, during growth on non-fermentable carbon sources (Kayingo et al., 2009). The data herein revealed that the Stl1 homologs (VDAG_00764 and VDAG_05113) were evidently downregulated in response to either of the two nitrogen sources, presumably in order to remodel the glycerol balance because intense metabolic activity may consume glycerol as carbon source, and further inferring a connection between nitrogen and carbon metabolism. GK (glycerol kinase) can phosphorylate glycerol to glycerol-3-phosphate (G3P) for glycerol assimilation (Wilk et al., 2020). TGL (triacylglycerol lipase) can gradually hydrolyze triglycerides into glycerol and fatty acids (Wadsater et al., 2014). Herein, glycerol was accumulated because of activated TGL during nitrate utilization, though both nitrogen assimilations caused downregulated Stl1 and GK.

The regulation of the expression of two bZIP transcription factors in response to the different nitrogen sources was

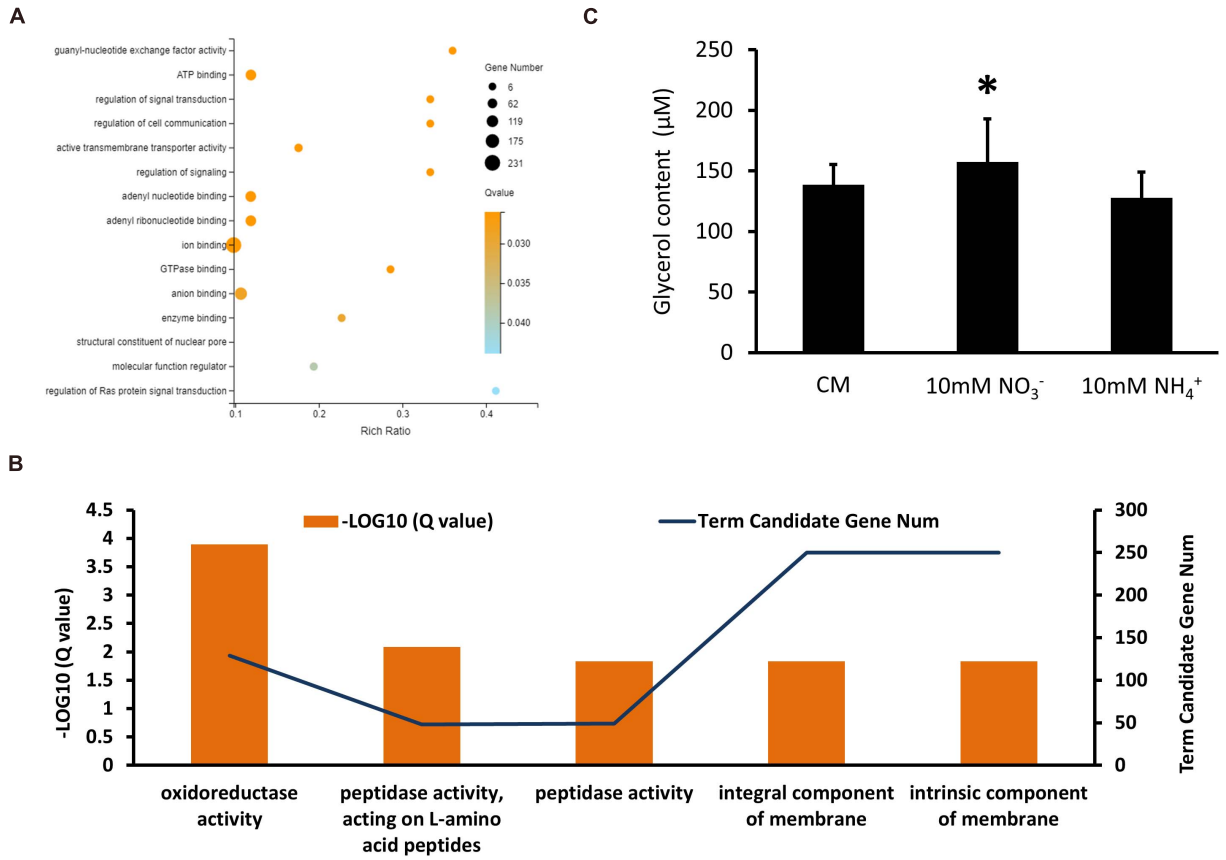


FIGURE 9 | Enrichment analyses of differentially expressed genes (DEGs) in *Verticillium dahliae* in response to ammonium sources. **(A)** A bubble chart showing DEGs upregulated during ammonium utilization and enriched in 13 GO terms. The size of the bubbles is proportional to the total number of DEGs. The *X*-axis indicates enrichment ratio and the *Y*-axis indicates the GO terms. **(B)** A linear graph showing DEGs downregulated in ammonium utilization and enriched in five GO terms. **(C)** The bar chart showed glycerol content of the XS11 strain treated with 10 mM NO_3^- or 10 mM NH_4^+ for 24 h. Error bars represent the standard deviation based on three independent replicates. Asterisks indicate significant differences (*$P < 0.05$).

notable in these experiments. There was a striking upregulation of VDAG_08640, a bZIP transcription factor in response to both nitrogen sources. Previous work also revealed that this particular gene was differentially expressed during microsclerotia development, but does not regulate microsclerotia formation, the sensitivity to hydrogen peroxide, and virulence, hence was not considered to be indispensable in *V. dahliae* (Fang et al., 2017). Deletion of this gene caused an affect on nitrate utilization, suggesting a role of VDAG_08640 in nitrate metabolism. Nevertheless, involvement of the bZIP transcription factor in nitrogen metabolism is novel and requires further study. *VdHapX* is also a bZIP transcription factor contributing to adaptation to iron starvation and excesses, hydrogen peroxide detoxification, microsclerotium formation and virulence in *V. dahliae* (Wang et al., 2018). In *Fusarium graminearum*, the transcription factor cascade FgAreA-HapX link iron homeostasis to nitrogen metabolism: reduced activity of cytosolic Fe–S proteins nitrite reductase causes high expression of FgHapX via activating transcription factor FgAreA, then FgHapX suppresses another iron regulator FgSreA to maintain iron homeostasis (Wang Z. et al., 2019). Here, the expression of *VdHapX* was depressed in response both two nitrogen assimilation, while its knockout led to a clear growth defect during nitrate utilization. Potentially, deletion of *VdHapX* affected iron homeostasis, reducing the activity of nitrite reductase which is a key enzyme in nitrate assimilation and finally caused decreased growth during utilizing nitrate. However, nitrite reductase is dispensable in ammonium assimilation, thus deletion of *VdHapX* failed to impact this process. Its decreased expression may be related to *V. dahliae* placing increased energy into nitrogen utilization and depression of other metabolic processes for survival in a changing environment.

Overlap Between Responses to Different Inorganic Nitrogen Sources in *V. dahliae*

Although there were various differences in DEGs associated with different metabolic pathways when supplied with preferred and alternative nitrogen sources in *V. dahliae*, some regulatory mechanisms are similar in both metabolic processes. The cell cycle, involving a series of complex cellular events, is vital to

morphological changes related to infectious growth or infection-related morphogenesis in some fungal pathogens (Sendinc et al., 2015; Jiang et al., 2016). The cell cycle machinery and metabolism are interconnected. The initiation of cell cycle critically depends on the availability of metabolites, and it maintains cell survival and proliferation by regulating metabolic networks (Kaplon et al., 2015). In the absence of glucose, for example, cells may be arrested at the G1/S phase transition, demonstrating how carbon metabolism impacts cell cycle (Pardee, 1974). Herein, genes associated with cell cycle such as *VdMcm1* were downregulated during ammonium assimilation, inferring a link between nitrogen metabolism and the cell cycle. Deletion of *VdMcm1* resulted in significant reduction of hyphal growth (Xiong et al., 2016), which may explain the reduced growth when cultured on ammonium medium. The expression of VDAG_08640 was obviously reduced in the *VdMcm1* deletion mutant (Xiong et al., 2016), which inferred that VdMcm1 regulates nitrate utilization via affecting expression of VDAG_08640.

Expression of the gene encoding D-3-phosphoglycerate dehydrogenase (PGDH) (VDAG_01596), the committal enzyme in L-serine biosynthesis (Grant, 2018), was observably reduced in both conditions in which nitrate and ammonium were independently supplied as nitrogen sources. L-Serine is a crucial amino acid because it serves as a precursor the production of vital metabolites. The most-well-studied PGDHs are bacterial, primarily from *Escherichia coli* and *Mycobacterium tuberculosis*, showing the catalytic activity of PGDH and L-serine form a feedback loop and it is involved in bacterial motility, adherence and virulence (Zhao and Winkler, 1996; Peters-Wendisch et al., 2002; Yasuda et al., 2017). In plants, PGDH plays a vital function in plant development and metabolism, and participates in some pathways linking carbon and nitrogen metabolism and in maintaining cellular redox and energy levels in stress conditions (Igamberdiev and Kleczkowski, 2018). The ABC (ATP-binding cassette) transporter family, involved in resistance to azole antifungal agents, has been extensively documented (Coste et al., 2004). In *C. albicans*, *Cdr1* (*Candida* drug resistance) from the family of ABC transporters, a gene most often associated with an energy-dependent multidrug efflux pump, is the principal mediator of resistance to azoles due to transport phenomena (Sanglard et al., 1995). The efflux activity of *Cdr1* pumps is energy-dependent (Thomas et al., 2013). In our study, *Cdr1* homolog (VDAG_05293) expression was remarkably reduced when *V. dahliae* assimilated both nitrate and ammonium sources, which may indicate that the pathogen preserve energy to metabolize either of the supplied nitrogen sources.

CONCLUSION

Based on the results presented herein, it is clear that metabolism of different nitrogen sources in *V. dahliae* triggers major differences in gene transcription. In this study, we highlighted the convergent and distinctive regulatory mechanisms between preferred and alternative nitrogen metabolism in *V. dahliae*. RNA-Seq analyses suggested that assimilation and metabolism of both nitrogen sources intensively impact many aspects of the oxidoreductase activity, transporter activity, and metabolic processes. However, the expression analyses also indicate that nitrate and ammonium metabolism are under different regulatory control mechanisms even though there were clearly also increases in transcription and translation processes when either was used as the sole nitrogen source. During utilization of nitrate, the expression analyses suggest that ribosome biosynthesis and protein translation are activated but signal transduction and communication are reduced. Additionally, gene expression associated with cellular nitrogen compound biosynthetic and metabolic processes increased in response to nitrate rather than ammonium as the sole nitrogen source, and therefore preferred nitrogen metabolism is apparently more efficient and less energy intensive. Further, according to our results, *V. dahliae* inhibits some cellular processes such as the cell cycle as well as other metabolic process during assimilation and metabolism of ammonium. Glycerol and protein content measurements verified some of these changes in this study, and bZIP transcription factors play an important role in nitrogen metabolism. This study provides useful information into our understanding of how *V. dahliae* utilizes two different nitrogen sources which may provide insights into alternative disease control measures.

DATA AVAILABILITY STATEMENT

Raw Illumina sequences were deposited in the National Center for Biotechnology Information (NCBI) and can be accessed in the Short Read Archive (SRA) database (http://trace.ncbi.nlm.nih.gov/Traces/sra/) under accession SRX11048827–SRX11048835 for XS11, XS11- NO_3^-, and XS11-$NH4^+$ respectively.

AUTHOR CONTRIBUTIONS

YW and CT conceived the experiments. CT and WL performed the experiments. CT analyzed the data. CT, SK, and YW wrote the manuscript. All authors read and approved the manuscript.

FUNDING

The research was supported the National Natural Science Foundation of China (31971657).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.712701/full#supplementary-material

Supplementary Table 1 | Enrichment terms of significantly differentially expressed genes (corrected *P*-value < 0.01).

Supplementary Table 2 | Enrichment terms of significantly differentially expressed genes (corrected *P*-value < 0.05).

REFERENCES

Bai, C., Tesker, M., and Engelberg, D. (2015). The yeast Hot1 transcription factor is critical for activating a single target gene. STL1. *Mol. Biol. Cell* 26, 2357–2374. doi: 10.1091/mbc.e14-12-1626

Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29, 1165–1188.

Berne, S., and Javornik, B. (2016). "Signalling crosstalk of plant defence responses to Xylem-invading pathogens," in *Abiotic and Biotic Stress in Plants-Recent Advances and Future Perspectives*, eds A. Shanker and C. Shanker (Germany: Books on Demand).

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Coste, A. T., Karababa, M., Ischer, F., Bille, J., and Sanglard, D. (2004). TAC1, transcriptional activator of CDR genes, is a new transcription factor involved in the regulation of *Candida albicans* ABC transporters CDR1 and CDR2. *Eukaryot. Cell* 3, 1639–1652. doi: 10.1128/ec.3.6.1639-1652.2004

Deng, Y., Yang, F., and Naqvi, N. I. (2015). "The role of nutrients in fungal development and pathogenesis," in *Fungal Biomolecules: Sources, Applications and Recent Developments*, eds V. K. Gupta and R. L. Mach (Hoboken, NJ: John Wiley & Sons).

Divon, H. H., Rothan-Denoyes, B., Davydov, O., Di Pietro, A., and Fluhr, R. (2005). Nitrogen-responsive genes are differentially regulated in planta during *Fusarium oxyspsorum* f. sp. lycopersici infection. *Mol. Plant Pathol.* 6, 459–470. doi: 10.1111/j.1364-3703.2005.00297.x

Fang, Y., Xiong, D., Tian, L., Tang, C., Wang, Y., and Tian, C. (2017). Functional characterization of two bZIP transcription factors in *Verticillium dahliae*. *Gene* 626, 386–394. doi: 10.1016/j.gene.2017.05.061

Fernandez, J., and Wilson, R. (2012). Why no feeding frenzy? mechanisms of nutrient acquisition and utilization during infection by the rice blast fungus *Magnaporthe oryzae*. *Mol. Plant Microbe* 25, 1286–1293. doi: 10.1094/mpmi-12-11-0326

Fernandez, J., Wright, J. D., Hartline, D., Quispe, C. F., Madayiputhiya, N., and Wilson, R. A. (2012). Principles of carbon catabolite repression in the rice blast fungus: Tps1, Nmr1-3, and a MATE–family pump regulate glucose metabolism during Infection. *PLoS Genet.* 8:e1002673. doi: 10.1371/journal.pgen.1002673

Ferreira, C., van Voorst, F., Martins, A., Neves, L., Oliveira, R., Kielland-Brandt, M. C., et al. (2005). A member of the sugar transporter family, Stl1p is the glycerol/H+ symporter in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 16, 2068–2076. doi: 10.1091/mbc.e04-10-0884

Fisher, D. (2000). "Long-distance transport," in *Biochemistry and Molecular Biology of Plants*, eds B. B. Buchanan, W. Gruissem, and R. L. Jones (Hoboken, NJ: John Wiley & Sons).

Froeliger, E. H., Carpenter, B. E., and Froeliger, E. (1996). NUT1, a major nitrogen regulatory gene in *Magnaporthe grisea*, is dispensable for pathogenicity. *Mol. Gen. Genet.* 251, 647–656. doi: 10.1007/bf02174113

Gouzy, A., Poquet, Y., and Neyrolles, O. (2014). Nitrogen metabolism in *Mycobacterium tuberculosis* physiology and virulence. *Nat. Rev. Microbiol.* 12, 729–737. doi: 10.1038/nrmicro3349

Grant, G. A. (2018). D-3-Phosphoglycerate dehydrogenase. *Front. Mol. Biosci.* 5:110.

Gupta, K. J., Brotman, Y., Segu, S., Zeier, T., Zeier, J., Persijn, S. T., et al. (2013). The form of nitrogen nutrition affects resistance against *Pseudomonas syringae* pv. phaseolicola in tobacco. *J. Exp. Bot.* 64, 553–568. doi: 10.1093/jxb/ers348

Hannan, R. D., Drygin, D., and Pearson, R. B. (2013). Targeting RNA polymerase I transcription and the nucleolus for cancer therapy. *Expert Opin. Ther. Targets* 17, 873–878. doi: 10.1517/14728222.2013.818658

Hu, D., Xue, S., Zhao, C., Wei, M., Yan, H., Quan, Y., et al. (2018). Comprehensive profiling of lysine acetylome in baculovirus infected silkworm (*Bombyx mori*) cells. *Proteomics* 18:1700133. doi: 10.1002/pmic.201700133

Igamberdiev, A. U., and Kleczkowski, L. A. (2018). The glycerate and phosphorylated pathways of serine synthesis in plants: the branches of plant glycolysis linking carbon and nitrogen metabolism. *Front. Plant Sci.* 9:318.

James, M. J., and Zomerdijk, J. C. (2004). Phosphatidylinositol 3-kinase and mTOR signaling pathways regulate RNA polymerase I transcription in response to IGF-1 and nutrients. *J. Biol. Chem.* 279, 8911–8918. doi: 10.1074/jbc.m307735200

Ji, H., Lu, X., Zong, H., and Zhuge, B. (2018). Functional and expression studies of two novel STL1 genes of the osmotolerant and glycerol utilization yeast *Candida glycerinogenes*. *J. Gen. Appl. Microbiol.* 64, 121–126. doi: 10.2323/jgam.2017.10.001

Jiang, C., Xu, J.-R., and Liu, H. (2016). Distinct cell cycle regulation during saprophytic and pathogenic growth in fungal pathogens. *Curr. Genet.* 62, 185–189. doi: 10.1007/s00294-015-0515-9

Kaplon, J., van Dam, L., and Peeper, D. (2015). Two-way communication between the metabolic and cell cycle machineries: the molecular basis. *Cell Cycle* 14, 2022–2032. doi: 10.1080/15384101.2015.1044172

Kayingo, G., Bill, R., Calamita, G., Hohmann, S., and Prior, B. (2001). Microbial water and glycerol channels. *Curr. Top. Membr.* 51, 335–370. doi: 10.1016/s1063-5823(01)51010-8

Kayingo, G., Martins, A., Andrie, R., Neves, L., Lucas, C. N., and Wong, B. (2009). A permease encoded by STL1 is required for active glycerol uptake by *Candida albicans*. *Microbiology* 155(Pt 5), 1547–1557. doi: 10.1099/mic.0.023457-0

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Klosterman, S. J., Atallah, Z. K., Vallad, G. E., and Subbarao, K. V. (2009). Diversity, pathogenicity, and management of Verticillium species. *Annu. Rev. Phytopathol.* 47, 39–62. doi: 10.1146/annurev-phyto-080508-081748

Klosterman, S. J., Subbarao, K. V., Kang, S., Veronese, P., Gold, S. E., Thomma, B. P., et al. (2011). Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. *PLoS Pathog.* 7:e1002137. doi: 10.1371/journal.ppat.1002137

Kusano, M., Fukushima, A., Redestig, H., and Saito, K. (2011). Metabolomic approaches toward understanding nitrogen metabolism in plants. *J. Exp. Bot.* 62, 1439–1453. doi: 10.1093/jxb/erq417

Kusnadi, E. P., Hannan, K. M., Hicks, R. J., Hannan, R. D., Pearson, R. B., and Kang, J. (2015). Regulation of rDNA transcription in response to growth factors, nutrients and energy. *Gene* 556, 27–34. doi: 10.1016/j.gene.2014.11.010

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323.

Lippman, S. I., and Broach, J. R. (2009). Protein kinase A and TORC1 activate genes for ribosomal biogenesis by inactivating repressors encoded by Dot6 and its homolog Tod6. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19928–19933. doi: 10.1073/pnas.0907027106

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 1–21.

Marroquin-Guzman, M., and Wilson, R. A. (2015). GATA-dependent glutaminolysis drives appressorium formation in *Magnaporthe oryzae* by suppressing TOR inhibition of cAMP/PKA signaling. *PLoS Pathog.* 11:e1004851. doi: 10.1371/journal.ppat.1004851

Pardee, A. B. (1974). A restriction point for control of normal animal cell proliferation. *Proc. Natl. Acad. Sci. U. S. A.* 71, 1286–1290. doi: 10.1073/pnas.71.4.1286

Peters-Wendisch, P., Netzer, R., Eggeling, L., and Sahm, H. (2002). 3-Phosphoglycerate dehydrogenase from *Corynebacterium glutamicum*: the C-terminal domain is not essential for activity but is required for inhibition by L-serine. *Appl. Microbiol. Biot.* 60, 437–441. doi: 10.1007/s00253-002-1161-y

Randhawa, A., Kundu, D., Sharma, A., Prasad, R., and Mondal, A. K. (2019). Overexpression of the CORVET complex alleviates the fungicidal effects of fludioxonil on the yeast *Saccharomyces cerevisiae* expressing hybrid histidine kinase 3. *J. Biol. Chem.* 294, 461–475. doi: 10.1074/jbc.ra118.004736

Ries, L. N. A., Beattie, S., Cramer, R. A., and Goldman, G. H. (2018). Overview of carbon and nitrogen catabolite metabolism in the virulence of human pathogenic fungi. *Mol. Microbiol.* 107, 277–297. doi: 10.1111/mmi.13887

Sanglard, D., Kuchler, K., Ischer, F., Pagani, J., Monod, M., and Bille, J. (1995). Mechanisms of resistance to azole antifungal agents in *Candida albicans* isolates from AIDS patients involve specific multidrug transporters. *Antimicrob. Agents Chemother.* 39, 2378–2386. doi: 10.1128/aac.39.11.2378

Sendinc, E., Jambhekar, A., and Shi, Y. (2015). Remodeling your way out of cell cycle. *Cell* 162, 237–238. doi: 10.1016/j.cell.2015.06.050

Singh, S., Braus-Stromeyer, S. A., Timpner, C., Lohaus, G., Reusche, M., Knüfer, J., et al. (2010). Silencing of Vlaro 2 for chorismate synthase revealed that the phytopathogen *Verticillium longisporum* induces the cross-pathway control in the xylem. *Appl. Microbiol. Biot.* 85, 1961–1976. doi: 10.1007/s00253-009-2269-0

Slavov, N., Semrau, S., Airoldi, E., Budnik, B., and van Oudenaarden, A. (2015). Differential stoichiometry among core ribosomal proteins. *Cell Rep.* 13, 865–873. doi: 10.1016/j.celrep.2015.09.056

Snoeijers, S. S., Pérez-García, A., Joosten, M. H., and De Wit, P. J. (2000). The effect of nitrogen on disease development and gene expression in bacterial and fungal plant pathogens. *Eur. J. Plant Pathol.* 106, 493–506.

Spriggs, K. A., Bushell, M., and Willis, A. E. (2010). Translational regulation of gene expression during conditions of cell stress. *Mol. Cell* 40, 228–237. doi: 10.1016/j.molcel.2010.09.028

Tang, C., Jin, X., Klosterman, S. J., and Wang, Y. (2020a). Convergent and distinctive functions of transcription factors VdYap1, VdAtf1, and VdSkn7 in the regulation of nitrosative stress resistance, microsclerotia formation, and virulence in *Verticillium dahliae*. *Mol. Plant Pathol.* 21, 1451–1466.

Tang, C., Li, T., Klosterman, S. J., Tian, C., and Wang, Y. (2020b). The bZIP transcription factor VdAtf1 regulates virulence by mediating nitrogen metabolism in *Verticillium dahliae*. *New Phytol.* 226, 1461–1479. doi: 10.1111/nph.16481

Thalineau, E., Truong, H.-N., Berger, A., Fournier, C., Boscari, A., Wendehenne, D., et al. (2016). Cross-regulation between N metabolism and nitric oxide (NO) signaling during plant immunity. *Front. Plant Sci.* 7:472.

Thomas, E., Roman, E., Claypool, S., Manzoor, N., Pla, J., and Panwar, S. L. (2013). Mitochondria influence CDR1 efflux pump activity, Hog1-mediated oxidative stress pathway, iron homeostasis, and ergosterol levels in *Candida albicans*. *Antimicrob. Agents Chemother.* 57, 5580–5599. doi: 10.1128/aac.00889-13

Timpner, C., Braus-Stromeyer, S. A., Tran, V. T., and Braus, G. H. (2013). The Cpc1 regulator of the cross-pathway control of amino acid biosynthesis is required for pathogenicity of the vascular pathogen *Verticillium longisporum*. *Mol. Plant Microbe Interact.* 26, 1312–1324. doi: 10.1094/mpmi-06-13-0181-r

Tudzynski, B. (2014). Nitrogen regulation of fungal secondary metabolism in fungi. *Front. Microbiol.* 5:656.

Wadsater, M., Barauskas, J., Nylander, T., and Tiberg, F. (2014). Formation of highly structured cubic micellar lipid nanoparticles of soy phosphatidylcholine and glycerol dioleate and their degradation by triacylglycerol lipase. *ACS Appl. Mater. Inter.* 6, 7063–7069. doi: 10.1021/am501489e

Wang, M., Gu, Z., Wang, R., Guo, J., Ling, N., Firbank, L. G., et al. (2019). Plant primary metabolism regulated by nitrogen contributes to plant–pathogen interactions. *Plant Cell Physiol.* 60, 329–342. doi: 10.1093/pcp/pcy211

Wang, Y., Deng, C., Tian, L., Xiong, D., Tian, C., and Klosterman, S. J. (2018). The transcription factor VdHapX controls iron homeostasis and is crucial for virulence in the vascular pathogen *Verticillium dahliae*. *MSphere* 3, e400–e418.

Wang, Y., Xiao, S., Xiong, D., and Tian, C. (2013). Genetic transformation, infection process and qPCR quantification of *Verticillium dahliae* on smoke-tree *Cotinus coggygria*. *Australas. Plant Pathol.* 42, 33–41. doi: 10.1007/s13313-012-0172-0

Wang, Z., Ma, T., Huang, Y., Wang, J., Chen, Y., Kistler, H. C., et al. (2019). A fungal ABC transporter FgAtm1 regulates iron homeostasis via the transcription factor cascade FgAreA-HapX. *PLoS Pathog.* 15:e1007791. doi: 10.1371/journal.ppat.1007791

Wilk, P., Kuśka, K., Wątor, E., Małecki, P. H., Woś, K., Tokarz, P., et al. (2020). Structural characterization of Glycerol kinase from the thermophilic fungus *Chaetomium thermophilum*. *Int. J. Mol. Sci.* 21:9570. doi: 10.3390/ijms21249570

Wilson, R. A., Gibson, R. P., Quispe, C. F., Littlechild, J. A., and Talbot, N. J. (2010). An NADPH-dependent genetic switch regulates plant infection by the rice blast fungus. *Proc. Natl. Acad. Sci. U. S. A.* 107, 21902–21907. doi: 10.1073/pnas.1006839107

Xiong, D., Wang, Y., Tian, L., and Tian, C. (2016). MADS-Box transcription factor VdMcm1 regulates conidiation, microsclerotia formation, pathogenicity, and secondary metabolism of *Verticillium dahliae*. *Front. Microbiol.* 7:1192.

Yadeta, K., and Thomma, B. (2013). The xylem as battleground for plant hosts and vascular wilt pathogens. *Front. Plant Sci.* 4:97.

Yasuda, M., Nagata, S., Yamane, S., Kunikata, C., Kida, Y., Kuwano, K., et al. (2017). *Pseudomonas aeruginosa* serA gene is required for bacterial translocation through Caco-2 cell monolayers. *PLoS One* 12:e0169367. doi: 10.1371/journal.pone.0169367

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14.

Yu, R., Campbell, K., Pereira, R., Björkeroth, J., Qi, Q., Vorontsov, E., et al. (2020). Nitrogen limitation reveals large reserves in metabolic and translational capacities of yeast. *Nat. Commun.* 11:1881.

Zhao, G., and Winkler, M. E. (1996). A novel alpha-ketoglutarate reductase activity of the serA-encoded 3-phosphoglycerate dehydrogenase of *Escherichia coli* K-12 and its possible implications for human 2-hydroxyglutaric aciduria. *J. Bacteriol.* 178, 232–239. doi: 10.1128/jb.178.1.232-239.1996

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 31 August 2021
doi: 10.3389/fmicb.2021.714651

Uncovering Diagnostic Value of Mitogenome for Identification of Cryptic Species *Fusarium graminearum Sensu Stricto*

Joanna Wyrębek[1], Tomasz Molcan[2], Kamil Myszczyński[3], Anne D. van Diepeningen[4], Alexander A. Stakheev[5], Maciej Żelechowski[1], Katarzyna Bilska[1] and Tomasz Kulik[1]**

[1]*Department of Botany and Nature Protection, University of Warmia and Mazury in Olsztyn, Olsztyn, Poland,* [2]*Department of Bioinformatics, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland,* [3]*Molecular Biology Laboratory, Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, Olsztyn, Poland,* [4]*Biointeractions & Plant Health, Wageningen Plant Research, Wageningen, Netherlands,* [5]*Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia*

OPEN ACCESS

Edited by:
Vaibhav Srivastava, Royal Institute of Technology, Sweden

Reviewed by:
Kerry O'Donnell, National Center for Agricultural Utilization Research (NCAUR), United States
John Carlyle Kennell, Saint Louis University, United States

***Correspondence:**
Joanna Wyrębek
joanna.wyrebek@uwm.edu.pl
Tomasz Kulik
tomaszkulik76@gmail.com

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *27 May 2021*
Accepted: *20 July 2021*
Published: *31 August 2021*

Citation:
Wyrębek J, Molcan T, Myszczyński K, van Diepeningen AD, Stakheev AA, Żelechowski M, Bilska K and Kulik T (2021) Uncovering Diagnostic Value of Mitogenome for Identification of Cryptic Species Fusarium graminearum Sensu Stricto. Front. Microbiol. 12:714651. doi: 10.3389/fmicb.2021.714651

Fungal complexes are often composed of morphologically nearly indistinguishable species with high genetic similarity. However, despite their close relationship, they can exhibit distinct phenotypic differences in pathogenicity and production of mycotoxins. Many plant pathogenic and toxigenic fungi have been shown to consist of such cryptic species. Identification of cryptic species in economically important pathogens has added value in epidemiologic studies and provides opportunities for better control. Analysis of mitochondrial genomes or *mitogenomics* opens up dimensions for improved diagnostics of fungi, especially when efficient recovery of DNA is problematic. In comparison to nuclear DNA, mitochondrial DNA (mtDNA) can be amplified with improved efficacy due to its multi-copy nature. However, to date, only a few studies have demonstrated the usefulness of mtDNA for identification of cryptic species within fungal complexes. In this study, we explored the value of mtDNA for identification of one of the most important cereal pathogens *Fusarium graminearum sensu stricto* (*F.g.*). We found that homing endonucleases (HEGs), which are widely distributed in mitogenomes of fungi, display small indel polymorphism, proven to be potentially species specific. The resulting small differences in their lengths may facilitate further differentiation of *F.g.* from the other cryptic species belonging to *F. graminearum* species complex. We also explored the value of SNP analysis of the mitogenome for typing *F.g.* The success in identifying *F.g.* strains was estimated at 96%, making this tool an attractive complement to other techniques for identification of *F.g.*

Keywords: ***Fusarium graminearum sensu stricto*, *Fusarium graminearum* species complex, homing endonucleases, introns, mobile genetic elements, identification**

INTRODUCTION

Fusarium graminearum sensu stricto (*F.g.*) ranks number 4 in the top 10 most important and best-studied species that cause diseases on agriculturally important plants (Dean et al., 2012). The fungus is often involved in two diseases of cereals: fusarium head blight (FHB) of wheat and barley and fusarium ear rot (FER) of maize, which lead to major losses for grain production worldwide. Besides yield reduction, *F.g.* is able to produce mycotoxins, among which trichothecenes pose a serious hazard to human and animals (Dong et al., 2020; Yao et al., 2020). *F.g.* belongs to the monophyletic fungal complex referred to as *F. graminearum* species complex (FGSC). This complex includes 16 genetically characterized cryptic species (Sarver et al., 2011), several of which are involved in cereal diseases in certain agricultural areas. The species are difficult to identify by morphological characters and often share a high DNA sequence similarity (Walkowiak et al., 2016). However, despite their close relationship, FGSC species and even strains within species can exhibit distinct phenotypic differences in pathogenicity and mycotoxin production, while some strains lack pathogenicity on certain hosts (Jianbo, 2014; van der Lee et al., 2015; Walkowiak et al., 2016).

The increasing number of genomic sequences that are deposited in public databases has enabled research in genomic features that contribute to phenotypic variation and niche specialization within FGSC. Genome comparisons revealed significant divergence between them that can be mostly linked to single-nucleotide polymorphisms (SNPs), insertions/deletions (indels), and gene content variation, which is species specific or fixed in certain populations. Comparative genomics provides insights into the evolutionary processes contributing to pathogen divergence at both the macroevolutionary and microevolutionary scales (Walkowiak et al., 2016; Kelly and Ward, 2018).

Besides nuclear genome analyses, analysis of fusarium mitogenomes may provide useful data to study phylogenetic relationships and evolution. Despite high conservation, mitogenomes of fungi within complexes often exhibit a significant degree of polymorphism (Brankovics et al., 2017, 2020). This variation largely results from the irregular distribution of mobile genetic elements (MGEs) such as introns and associated homing endonucleases (HEGs; Basse, 2010; Joardar et al., 2012; Pogoda et al., 2019). MGEs comprise a substantial fraction of fungal mitogenomes and display a different evolutionary history from other mitochondrial loci (Guha et al., 2018; Kolesnikova et al., 2019). Their mosaic distribution throughout evolution can be explained with the "aenaon" model, which combines "intron-early" evolution enriched by "intron-late" events through recombination involving vertical and horizontal gene transfer (Megarioti and Kouvelis, 2020). Notably, variation in MGE content appears to be dependent on taxonomic sampling. This variation can be highly conserved, indicating a low mobility of ancestrally acquired MGEs (Kulik et al., 2020b). In contrast, more recent enrichment events result in increased mosaicism of MGE patterns that can be observed at the population level (Wolters et al., 2015).

Fungi of the genus *Fusarium* display a high variation in MGE content, from MGE-poor to MGE-rich mitogenomes (Losada et al., 2014; Brankovics et al., 2017). Our previous study performed on a large collection of *F.g.* strains has shown that the mitogenome of this fungus is MGE rich and mainly includes introns from the group I intron family, which harbor either LAGLIDADG or GIY-YIG HEGs. Comparison of mitogenomes of geographically diverse *F.g.* strains did not reveal any population-specific profiles, thus supporting the hypothesis on ancestral acquisition of HEGs. Assuming their early acquisition and low mobility, recently diverging cryptic species may thus share a similar content of MGEs (Kulik et al., 2020b). This hypothesis will be tested using mitogenomes obtained in this study.

Variation in mitogenomes of closely related fungi may also be detected by phylogenetic conflict between single-gene trees. This phenomenon was discovered among and within members of closely related fungal complexes: the *Fusarium fujikuroi* species complex (FFSC) and *Fusarium oxysporum* species complex (FOSC; Brankovics et al., 2020). Conflicting tree topologies indicate incomplete lineage sorting from ancestral polymorphism or more recent interspecies gene flow since both can result in similar phylogenetic patterns (Fourie et al., 2013, 2018). The phylogenetic discordance complicates interpretation of phylogenetic reconstructions and largely limits clustering-based identification approaches (Wu et al., 2018).

Besides its evolutionary value, mitochondrial DNA (mtDNA) is promising in the diagnostics of fungi, especially with regard to its multi-copy nature, which facilitates its high recovery and amplification success (Santamaria et al., 2009). Both MGE content variation and SNP in mitogenomes could be used for identification purposes of fusaria (Wu et al., 2019; Kulik et al., 2020b). However, discriminating cryptic species based on mitochondrial sequences is challenging due to the high degree of conservation, and detecting species-specific sites requires screening of a large set of genomic data to differentiate between intraspecific and interspecific variation (Kulik et al., 2015, 2020a).

The analysis of SNPs provides an excellent tool for identifying and typing species (Uelze et al., 2020a). In general, SNPs display relatively low mutation rates and are evolutionarily stable (Leekitcharoenphon et al., 2012). Usually, SNPs are identified by mapping sequence data from individual strains against a closely related reference genome. This type of analysis relies on the generation of a certain number of SNPs. It includes all studied genomes, including the reference genome (Uelze et al., 2020b).

The aims of this study are (1) to explore whether MGE polymorphisms can be used for identification of *F.g.*, (2) discover the value of SNP analysis for typing *F.g.*, and (3) confirm the utility of mt-based phylogenetic approach for the determination of *F.g.* identity.

In total, 122 strains were analyzed: 88 strains of *F.g.* and 34 strains representing other members of the FGSC. We discovered that HEGs present in mitogenomes of the FGSC display small indel polymorphism, which facilitates recognition of *F.g.* We also showed that whole-mitogenome SNP analysis enables typing of *F.g.* with 96% confidence, which can make this tool a valuable complement to other diagnostic tools for *F.g.* A phylogenetic analyses based on core mitochondrial genes showed different topologies in the reconstructed phylogenetic trees and did not cluster strains of *F.g.* into single

species-specific clades, which neglects the value of this approach for determination of species identity.

MATERIALS AND METHODS

Fungal Strains

In total, 122 strains were analyzed in this study, with 88 *F.g.* strains and 34 strains representing all known cryptic species from the FGSC complex: *Fusarium acaciae-mearnsii* (three strains), *Fusarium aethiopicum* (two strains), *Fusarium asiaticum* (three strains), *Fusarium austroamericanum* (two strains), *Fusarium boothii* (three strains), *Fusarium brasilicum* (two strains), *Fusarium cortadariae* (three strains), *Fusarium gerlachii* (two strains), *Fusarium louisianense* (two strains), *Fusarium meridionale* (one strain), *F. meridionale* × *F. asiaticum* hybrid strain (one strain), *Fusarium mesoamericanum* (two strains), *Fusarium nepalense* (two strains), *Fusarium ussurianum* (two strains), *Fusarium vorosii* (three strains), and strain CBS 123663, which lacks a Latin binomial. All strains used in this study are included in **Supplementary File S1**.

DNA Extraction and Sequencing

DNA from fungal strains was extracted from 0.1 g of mycelium harvested from PDA medium with the use of the Quick-DNA Plant/Seed Miniprep Kit (Zymo Research, Irvine, CA, United States) and Genomic Mini AX Food (Gdynia, Poland) according to the manufacturer's protocols. Genome libraries were constructed using either a TruSeq DNA PCR-free library preparation kit (Illumina, San Diego, CA, United States) or a KAPA HyperPlus Kit (Roche Sequencing Solutions, Pleasanton, CA, United States). Whole-genome sequencing was performed by Macrogen (Seoul, South Korea). An Illumina HiSeq X 10 platform was used to sequence the genomes using a paired-end read length of 2 × 150 bp with an insert size of 350 bp.

Assembly and Annotation of Mitogenomes

To perform *de novo* assembly of fungal mitogenomes from whole-genome data, NOVOPlasty 2.7.2 (Dierckxsens et al., 2017) was used. HEGs were identified using NCBI's ORF Finder, Blast+ (v. 2.9.0; Johnson et al., 2008), and Geneious Prime software (Biomatters Ltd., New Zealand). Annotations were performed using MFannot, InterPro (Mitchell et al., 2015), CD-Search (Marchler-Bauer and Bryant, 2004), and Geneious Prime software (Biomatters Ltd., New Zealand). Annotation of tRNA genes was improved using tRNAscan-SE (Pavesi et al., 1994). Annotations were manually verified and corrected. Complete mitogenomes were deposited in the NCBI database under the GenBank accession numbers given in **Supplementary File S1**.

Exploration of MGE Diversity

A total of 122 mitogenomes were compared through multiple sequence alignments. The analysis was performed with progressive Mauve (Darling et al., 2010) implemented in Geneious Prime software to examine the distribution of MGEs in the mitogenomes.

SNP Analysis

SNP analysis was performed on the interactive web-based platform Edge bioinformatics.[1] SNPs were detected in multiple comparisons of assembled mitogenomes of the strains to the reference whole genome of PH-1 of *F.g.* (accession number: MH412632). Eight mitogenomes from closely related morphospecies (*Fusarium culmorum*, *Fusarium cerealis*, and *Fusarium pseudograminearum*) were also included in the analysis.

Phylogenic Analysis

Phylogenetic analyses were performed on single-gene alignments of coding sequences of 15 core mitochondrial genes (*cox 1–3*, *cob*, *atp6*, *atp8*, *atp9*, *nad1-6*, *nad4L*, and ribosomal *rnl* gene) from 130 strains: 122 mitogenomes of FGSC and eight mitogenomes from closely related morphospecies (*F. culmorum*, *F. cerealis*, and *F. pseudograminearum*). Coding sequence data (CDS) from 15 core mitochondrial genes was first aligned using MAFFT software (v.7.453; Katoh and Toh, 2010) with default settings. Single-gene phylogenetic analyses were performed for *atp9*, *cox1*, and *cox2*. The remaining single-gene alignments were excluded for further phylogenetic analyses due to low sequence polymorphism. The best partition schemes and corresponding substitution models were estimated using PartitionFinder2 (Lanfear et al., 2016). Afterwards, based on the alignment and obtained models, maximum likelihood analysis was conducted using IQ-TREE 2.0.3 with 1,000 ultrafast bootstrap (Minh et al., 2020). *F. pseudograminearum* was used as an outgroup.

To reveal variation in mitochondrial CDS, core mitochondrial genes were extracted and aligned separately using MAFFT software (v.7.453; Katoh and Standley, 2013). Gene polymorphism analyses were conducted for each CDS based on the alignment of 130 strains. Variation within each CDS was identified as a SNP or indel and counted with the use of an in-house Python script. Additionally, each SNP was characterized as synonymous or nonsynonymous. Nucleotide diversity values (π) for each CDS were calculated with TASSEL software (v.5.2.40; Bradbury et al., 2007). As nucleotide diversity is based only on nucleotide substitutions, the number of indels and percentage of polymorphic sites are given for each CDS. To reveal variation in intronic sequences, conserved introns (found in all studied strains) were extracted and analyzed as described for mitochondrial CDS.

RESULTS

General Characteristics of Fungal Mitogenomes

Mitogenomes of species within the FGSC are highly conserved in terms of a set of 15 protein-coding genes, two rRNA genes (*rns* and *rnl*), and 28 tRNA genes, which were localized in the same order and orientation. Their mitogenomes displayed a similar GC content between 31.3 and 32.1%. However, mitogenome comparisons showed considerable differences in their size ranging from the smallest 88.409 bp mitogenome of

[1]https://www.edgebioinformatics.org/

CBS 119183 *Fusarium cortaderiae* to the largest 106.714 bp in CBS 119177 *F. vorosii*. This variation was mainly associated with a mosaic distribution of MGEs, which is described in one of the following sections.

All FGSC strains contained a large open reading frame with unknown function (LV-uORF; Al-Reedy et al., 2012), which is in every strain located between the *rnl* and *nad2* genes. This LV-uORF differed in size from 4.536 to 6.273 bp due to indel polymorphism (**Figure 1**).

Mobile Genetic Elements and Their Distribution in FGSC

A total of 45 introns and 56 HEGs were identified in the set of mitochondrial protein-coding genes of the studied strains. Nineteen introns were highly conserved and present in all studied strains. More than half of the introns ($n = 24$) identified were located in the three subunits of the cytochrome c oxidase (*cox* genes): 14 in *cox1*, 5 in *cox2*, and 5 in *cox3*. Positions of introns were highly conserved among different strains. All introns belonged to the group I intron family except i1 *cob* found in three strains: *F.g.* (INRA-156 and 16-462-z) and *F. vorosii* (CBS 119177), which was classified as belonging to the group II intron. This intron encoded a protein of the reverse transcriptase family. Most of group I introns harbored HEGs, of either the LAGLIDADG or GIY-YIG type, which can be determined based on differences in conserved amino acid motifs. Among 56 distinct HEGs, 33 were assigned as LAGLIDADG endonucleases and 16 as GIY-YIG endonucleases, while seven HEGs could not be precisely predicted as LAGLIDADG or GIY-YIG.

Eight introns showed very high similarity with no size difference variation, while 11 introns displayed some size variation due to small indels mostly below 100 bp in size. Nine introns showed large size variation (the i4 *nad2* intron, the i1 and i2 *cox2* introns, the i2 and i5 *cob* introns, the i3 *cox1* intron, the i1 *atp6* intron, and the i1 and i3 *cox3* introns), mostly due to the irregular presence of HEGs. HEGs and/or introns that are absent in at least one studied strain are termed "optional" throughout this manuscript.

From the total 46 optional HEGs, 11 were very limited in distribution. A single strain of *F. mesoamericanum* (CBS 110252) harbored two optional HEGs (144 aa and 170 aa) in the i5 *nad2* intron. Due to loss of conserved amino acid motifs reflecting their functional differences, these two HEGs were not precisely determined to either LAGLIDADG or GIY-YIG endonucleases. Similarly, a single strain of *F. vorosii* (CBS 119177) contained an optional LAGLIDADG (352 aa) found in the i1 *atp6* intron. The presence of two optional GIY-YIG (314 aa and 315 aa) and one LAGLIDADG (196 aa) sequences in intron i9 of the *cob* gene was found in a single strain of *F. asiaticum* (CBS 110258). More frequent distributions were observed for the remaining MGEs found in the *cob* gene. An optional LAGLIDADG (474 aa) located in the i2 *cob* intron was shared by two (CBS 316.73 and CBS 119170) of the three strains of *F. boothii*. Three strains, CBS 123662 (*F. acaciae-mearnsii*), CBS 110244 (*F. austroamericanum*), and CBS 123754 (*F. ussurianum*), harbored an optional LAGLIDADG (234 aa) in the i8 *cob* intron. Two other HEGs found in the *cob* gene were distributed in FGSC at 12.5% frequency and included an optional GIY-YIG (127 aa) in the i2 *cob* intron of

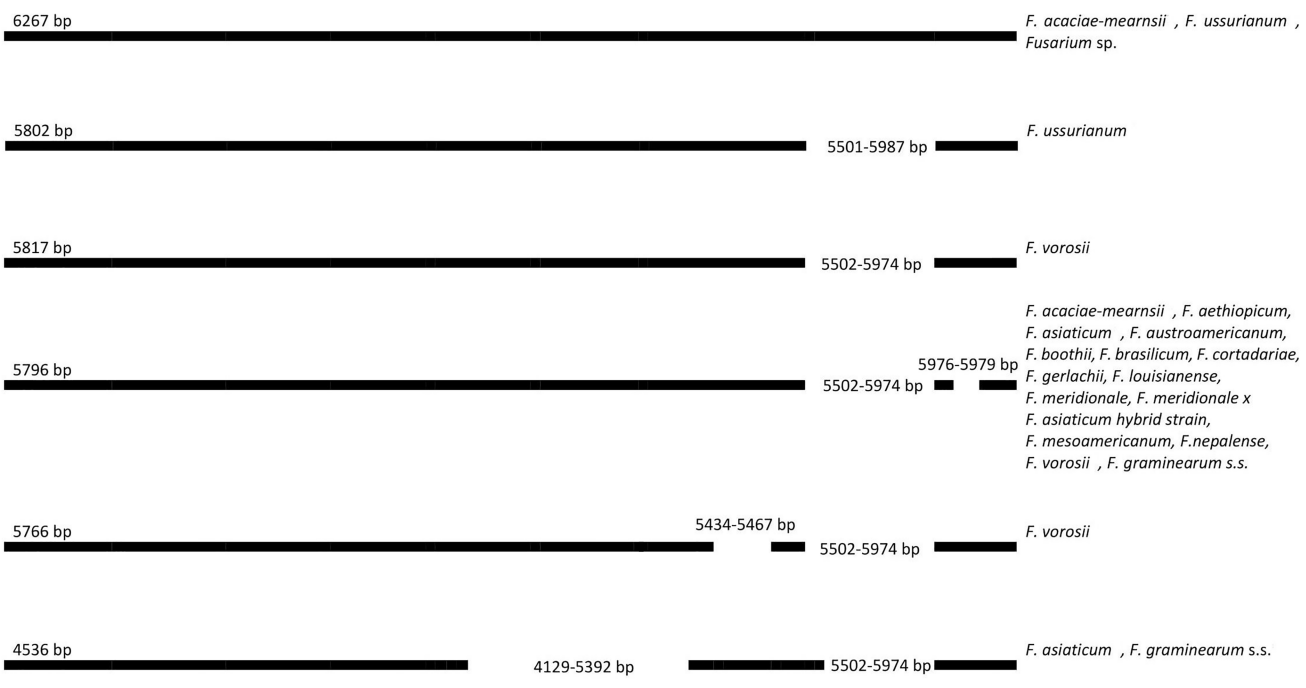


FIGURE 1 | Schematic position of indels in LV-uORF in 122 *Fusarium* strains.

F. brasilicum and *F. ussurianum* and GIY-YIG (314 aa) as well as LAGLIDADG (192 aa) in the i7 *cob* intron of all strains of *F. nepalense* and *F. vorosii*. The remaining 24 optional HEGs (**Supplementary File S1**) were more commonly distributed, ranging in frequency from 21.9 to 97.1% (**Supplementary File S1**; **Figures 2–6**).

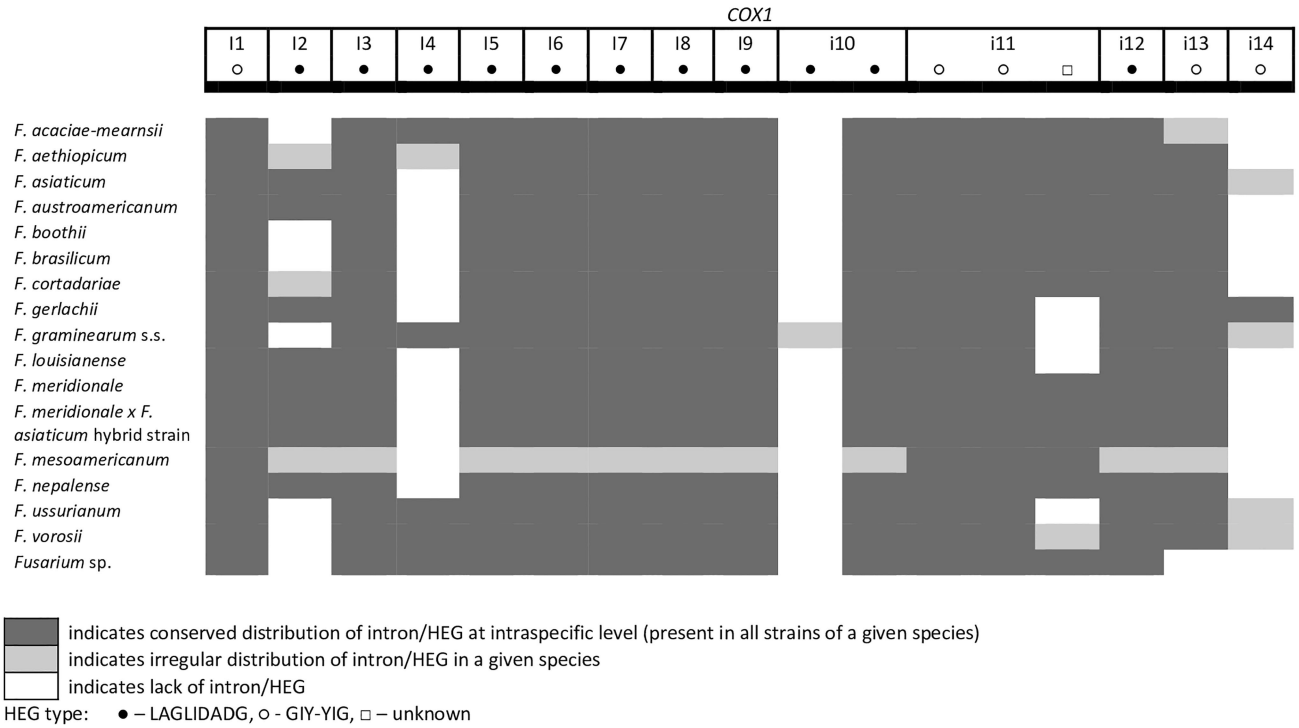


FIGURE 2 | Distribution of introns and associated HEGs found in the *COX1* gene.

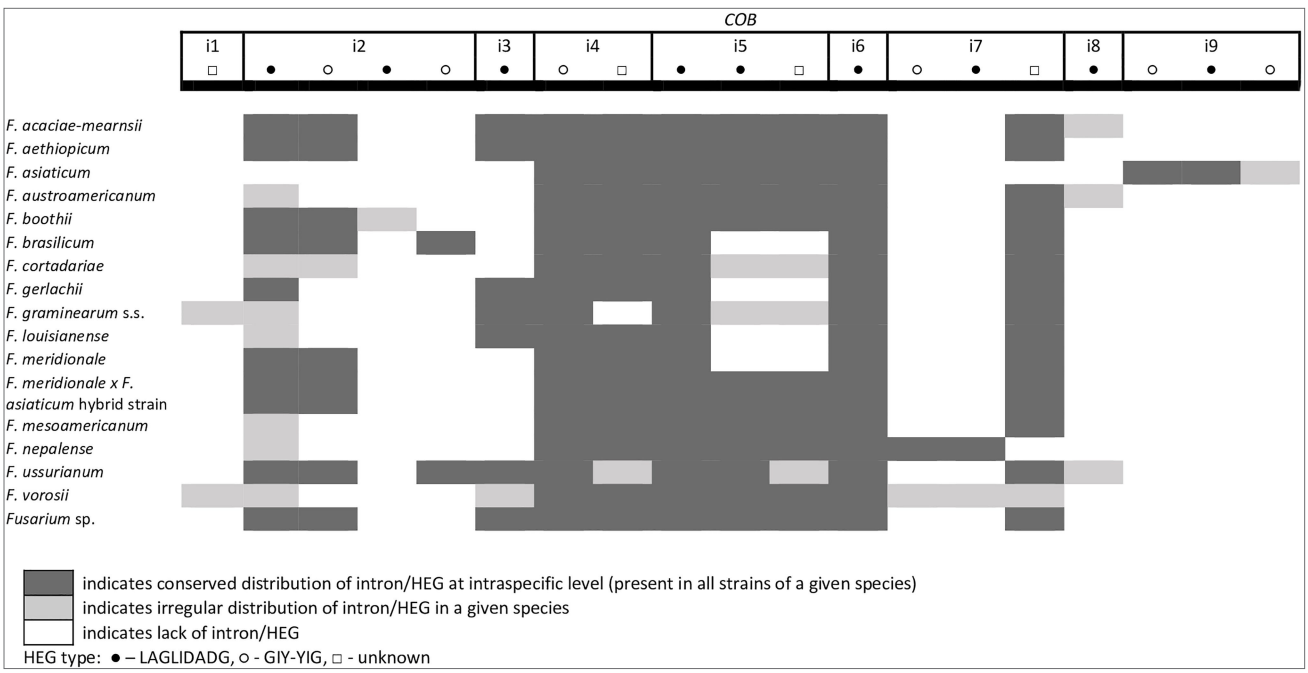


FIGURE 3 | Distribution of introns and associated HEGs found in the *COB* gene.

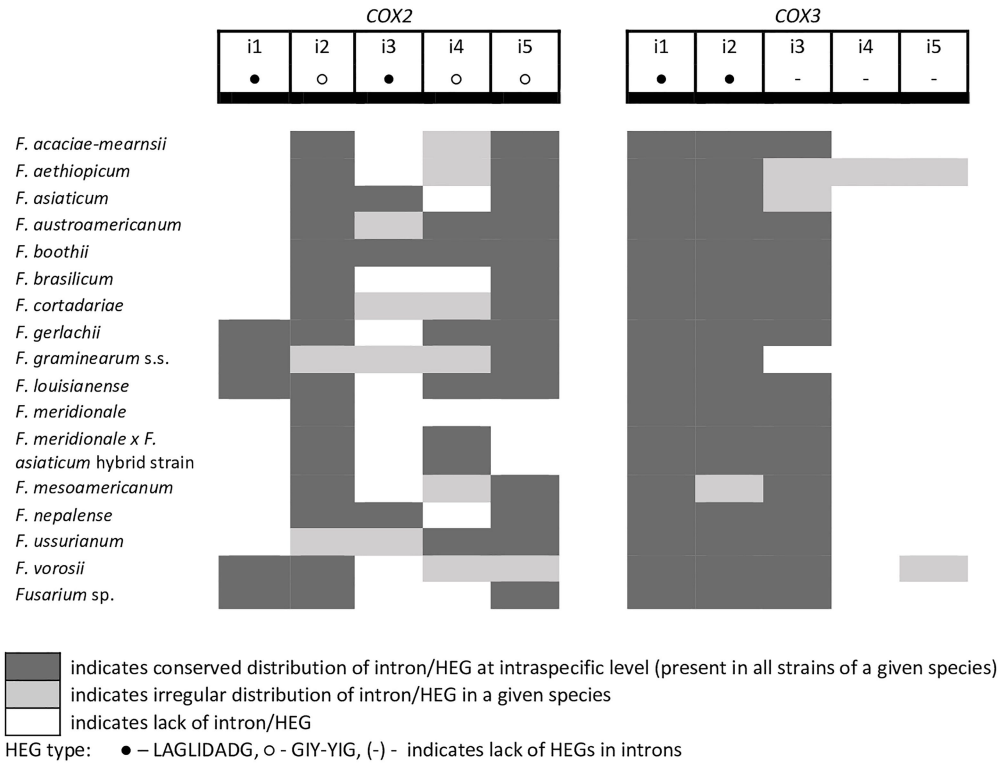


FIGURE 4 | Distribution of introns and associated HEGs found in the *COX2* and *COX3* genes.

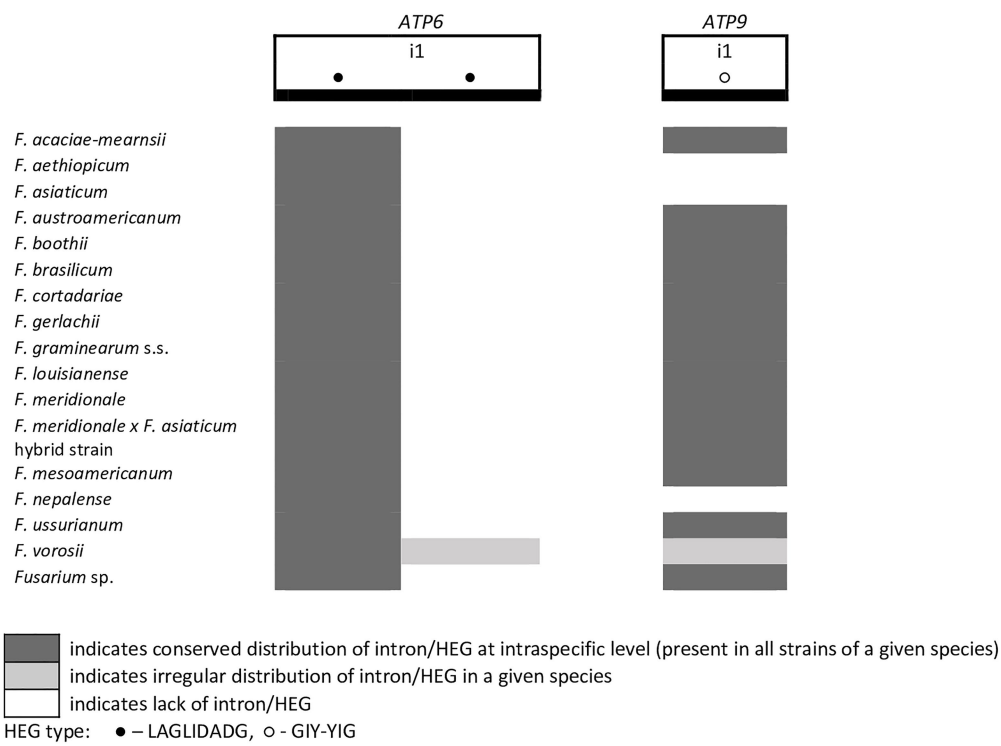


FIGURE 5 | Distribution of introns and associated HEGs found in the *ATP6* and *ATP9* genes.

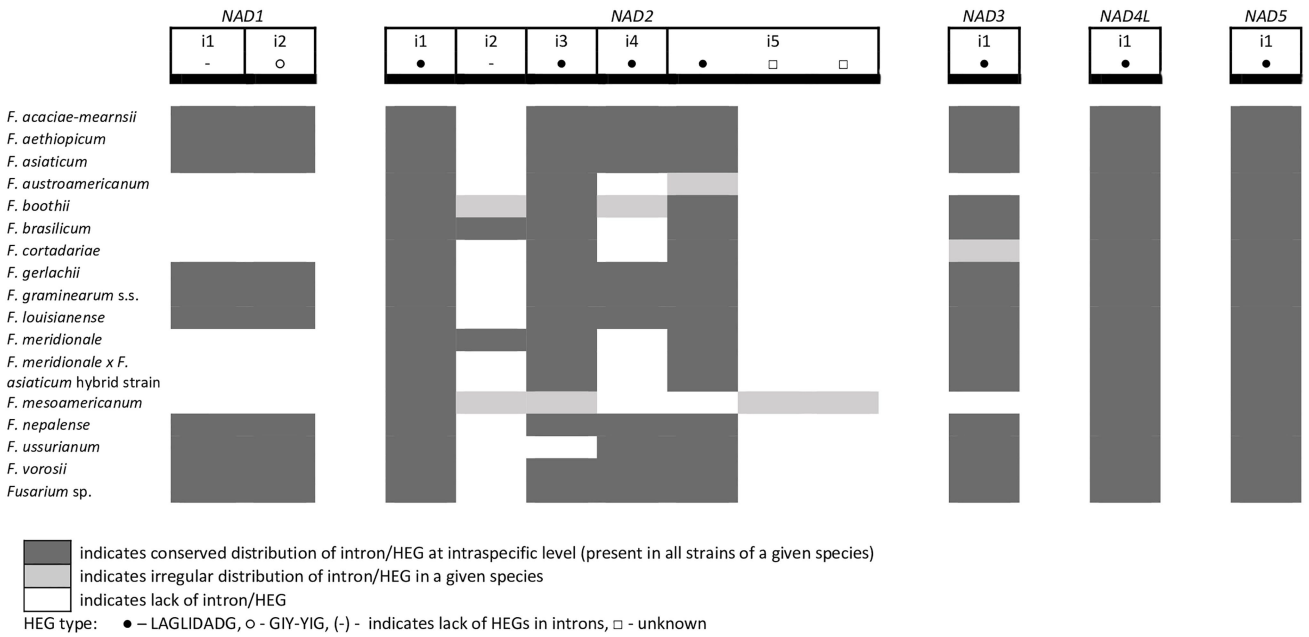


FIGURE 6 | Distribution of introns and associated HEGs found in the *NAD* genes.

A large collection of *F.g.* strains enabled the study of MGE distribution in their mitogenomes and possible cryptic divergence. Forty different MGEs were found in *F.g.*; however, none showed a species-specific conservation. We previously found that a unique MGE pattern in the *cox3* gene allows reliable differentiation of *F.g.* from other closely related morphospecies. This unique pattern mainly concerns the species-specific loss of the i3 intron, which was evident in all strains of *F.g.*, regardless of their geographic origin. In this study, we found that in contrast to *F.g.*, the i3 *cox3* intron was present in almost all other members of the FGSC, except for two strains of *F. asiaticum* (191,317 and 180,363; Yang et al., 2020; **Supplementary File S1**). However, besides their irregular distribution, HEGs can differ in size due to small indel polymorphism (**Table 1**). To explore the value of this type of polymorphism for identification of *F.g.*, we compared a total of 1,564 individual HEGs and found that the difference in HEG size observed in eight HEGs (**Table 1**) can be used as a unique diagnostic feature for *F.g.*

TABLE 1 | List of HEGs allowing differentiation of *F.g.s.s.* based on size difference.

Gene	Intron	HEG type	HEG size (nt)	
			F.g.s.s.	Other members of FGSC
nad5	i1 (IB)	LAGLIDADG	876	909
cox2	i5 (IC1)	GIY-YIG	1011	1242–1368
cob	i4 (ID)	GIY-YIG	642	870
cob	i7 (IC1)	unknown	1269	1419
cox1	i6 (IB)	LAGLIDADG	867	918
cox1	i7 (IB)	LAGLIDADG	957	972–1002
cox1	i8 (IB)	LAGLIDADG	1191	1221
atp6	i1 (IB)	LAGLIDADG	1035	1059

SNP Analysis

We performed a whole-mitogenome SNP analysis to reveal the utility of this type of approach for differentiation of *F.g.* from the other species. We calculated the number of SNPs for every mitogenome to reveal if interspecific variability in mtDNA is greater than intraspecific variability, indicating the existence of the so-called "barcoding gap" in the data set. SNPs were detected from the core genome of assembled mitogenomes by multiple comparisons of studied mitogenomes against the reference mitogenome of PH-1 strain (CBS 123657, accession number: MH412632). Analyses of variation in both CDS and conserved introns (**Tables 2** and **3**) showed that most SNPs could be found in intronic sequences. In general, SNP variability was low, ranging from 9 to 230 SNPs per strain (**Supplementary File S2**). In most cases, intraspecific variability did not exceed 30 SNPs. However, two exceptional strains of *F.g.*, INRA-156 and CBS 134070, yielded 32 and 36 SNPs, respectively, which exceeded the least interspecies variability (31 SNPs) found for *F. vorosii* strain CBS 119177. Thus, assuming intraspecific variability below 30 SNPs, two out of 88 strains of *F.g.* could not be determined based on SNP analysis, which indicates 96% accuracy of this SNP-based approach. Besides exceptional strain (CBS 119177), the least interspecific variability was found for two strains of *F. gerlachii*, yielding 36 SNPs each. Surprisingly, the highest number of interspecific SNPs was found for two strains of *F. ussurianum* (227 and 230 SNPs, respectively), which exceeded the number of SNPs found in case of three morphospecies: *F. cerealis*, *F. culmorum*, and *F. pseudograminearum*. This result may indicate that some cryptic species may have an accelerated

TABLE 2 | Variation in mitochondrial CDS among fungal strains.

Gene	Length (bp)	Synonymous SNPs	Nonsynonymous SNPs	Indels*	% PS	π
atp8	147	0	0	0	0	0
atp9	225	3	0	0	1.33	0.002
nad4L	270	1	0	0	0.37	0.00009
nad3	414	2	0	0	0.48	0.0005
nad6	681	0	0	0	0	0
cox2	750	5	0	0	0.67	0.0004
atp6	642–801	2	3	159	20.47	0.00324
cox3	810	25	8	0	4.20	0.00822
nad1	1110	0	0	0	0	0
cob	1173	2	0	0	0.17	0.00038
nad4	1455	2	0	0	0.14	0.00033
rps3	1482	1	2	0	0.20	0.00033
cox1	1593	31	0	0	1.95	0.00303
nad2	1665	2	1	0	0.18	0.00005
nad5	1989	1	1	0	0.10	0.00023

*% PS, percent of polymorphic sites and π, nucleotide diversity values. *Indels include single nucleotide insertions and deletions and deletion or duplication of longer tracts of DNA.*

TABLE 3 | Variation in conserved mitochondrial introns among fungal strains.

Intron name	Length (bp)	SNPs	Indels*	% PS	π
cob i4	2249	31	7	1.69	0.00193
cob i5	2467	16	1201	49.33	0.22581
cob i6	1061	7	0	0.66	0.00064
cox1 i1	1717	339	483	47.87	0.0293
cox1 i3	2398	449	1400	77.11	0.08436
cox1 i4	1375	9	0	0.65	0.00079
cox1 i5	1003	1	2	0.3	0.00027
cox1 i6	1154	6	45	4.42	0.00158
cox1 i7	1257	8	0	0.64	0.00026
cox1 i8	1325	476	275	56.68	0.00929
cox1 i9	2084	4	1078	51.92	0.00823
cox1 i10	2373	19	35	2.28	0.00595
cox1 i11	1198	1	0	0.08	0.00033
cox3 i1	1532	196	371	37.01	0.01505
cox3 i2	1875	682	499	62.99	0.01986
atp6 i1	3538	7	2086	59.16	0.00976
nad2 i1	1406	10	0	0.71	0.00073
nad4L i1	2008	34	208	12.05	0.01215
nad5 i1	1012	2	0	0.2	0.00014

*% PS, percent of polymorphic sites and π, nucleotide diversity values. *Indels include single nucleotide insertions and deletions and deletion or duplication of longer tracts of DNA.*

evolutionary rate with a unique evolutionary pattern in their mtDNA. In addition, the difference in evolutionary rate could be even observed at the strain level. Among the three strains of *F. vorosii*, CBS 119178 and CBS 123664 exhibited a similar number of interspecies SNPs (147 and 148 SNPs, respectively), which is in contrast to the exceptional strain (CBS 119177), which, as indicated above, yielded only 31 SNPs. An evidence for different evolutionary rates at the strain level was also observed by positioning some strains of the same species into scattered clades on a phylogenetic tree. This has been indicated in the next section.

Phylogenetic Analyses

Phylogenetic trees of *atp9*, *cox1*, and *cox2* genes possess different topologies, which can be explained by their different evolutionary histories. In addition, majority of the well-supported clades were multispecies. These results arise from low CDS variation in the studied strains, which were further confirmed by calculating nucleotide diversity values, number of polymorphic sites, and SNPs (**Table 2**). We also found that some strains of the same species were scattered in different clades of the tree. This was mostly evident in the *cox1* tree for strains from *F.g.* and other cryptic species such as *F. boothii*, *F. mesoamericanum*, *F. cortadariae*, and *F. vorosii* (**Supplementary File S4**), suggesting incomplete lineage sorting and introgression in the course of the evolution of the studied Fusaria.

DISCUSSION

The introduction of molecular tools has revolutionized the diagnostics of microorganisms, including fungal plant pathogens. The use of DNA-based markers opens possibilities for improved control of specific pathogens and pathotypes (Luchi et al., 2020). Identification of fungi based on mtDNA provides a valuable alternative to existing molecular approaches based on nuclear data, especially when efficient recovery of DNA is problematic. mtDNA can be amplified with higher success than nuclear DNA due to its high copy number in fungal cells (Krimitzas et al., 2013; Franco et al., 2017). However, exploration of mtDNA for diagnostic purposes may be challenging due to difficulties in detecting species-specific polymorphism. These difficulties are mostly reflected by (1) the highly conserved nature of mtDNA (Zhang et al., 2020), (2) intrastrain mosaicism in MGE patterns (Losada et al., 2014), and (3) phylogenetic discordance between different mitochondrial genes (Brankovics et al., 2017; Fourie et al., 2018). A previous comparative analysis of a large set of *F.g.* strains underlined the high conservation and similarity of mitogenomes in closely related morphospecies: *F. cerealis*, *F. culmorum*, and *F. pseudograminearum*. In general, mitogenomes of these closely related species were MGE rich, and their irregular distribution among strains determines, to a large extent, mitogenome variation. MGEs that do not display

mobility among other lineages can be fixed in species. Such targets are good candidates for developing diagnostic tools for identification of species; however, their identification requires comparison of a large set of target and nontarget strains. On the other hand, frequent mobility of MGEs could be incorporated into population genetic studies illustrating the local genetic variation and gene flow (Wolters et al., 2015; Wu et al., 2019).

In this study, we confirmed that the irregular distribution of MGEs does not follow divergence of cryptic species within the FGSC. This finding supports the hypothesis of low mobility of ancestrally acquired MGEs that are consequently shared by geographically diverse cryptic species (Kulik et al., 2020b). These ancestral elements may undergo changes throughout evolution, and assuming the lack of horizontal gene flow, these alternations should, at least in part, exhibit species-specific polymorphism. Indeed, in this study, we found that HEGs shared by different species can display small indel polymorphism that is species specific. The small differences in lengths of certain HEGs (**Table 1**) may facilitate further differentiation of *F.g.* from the other members of the FGSC.

In this study, we also explored the value of SNP analysis for typing *F.g.* This approach is increasingly used in the epidemiologic analyses of bacterial (Schürch et al., 2018) and fungal pathogens (Uelze et al., 2020b); however, to our knowledge, it has not yet been tested using fungal mtDNA. Successful identification of strains to the species level is achieved based on difference in intraspecific and interspecific SNP divergences (Bae et al., 2021). We found that intraspecific variation among *F.g.* strains was very low. Based on our assumption that intraspecific variation may not exceed 30 SNPs, the success rate in identifying strains of *F.g.* was estimated at 96%. The high confidence of this approach, its simplicity to perform, and the limited time needed for the analysis make this tool an attractive complement to other more refined techniques for identifying of *F.g.* (Wu et al., 2019; Kulik et al., 2020b).

Mitochondrial sequences can contribute to species identification due to their efficacy in revealing phylogenetic relationships among the studied taxa. However, we found that for *F.g.* and the other cryptic species within the FGSC, this approach is not discriminative enough mostly due to the highly conserved nature of mtDNA and lack of concordance between the different gene genealogies, which presumably results from incomplete ancestral lineage sorting. Our findings on the lack of concordance between different phylogenies of mitochondrial genes are in line with previous studies by Fourie et al. (2013, 2018) and Brankovics et al. (2020) who detected interspecies gene flow among mitogenomes of closely related members of the FFSC and FOSC. The success in detecting ancestral gene flow events through mitochondrial comparisons is due to the greater levels of conservation and synteny of fungal mitogenomes than observed in nuclear compartments (Brankovics et al., 2020). However, from a diagnostic point of view, this phylogenetic discordance largely limits clustering-based identification approaches (Wu et al., 2018).

Overall, the results presented in this study showed that the use of mtDNA provides valuable information for identification of cryptic species within fungal complexes. Although fungal mitogenomes lack a "universal barcode" for tagging cryptic species, they might display other patterns of species-specific conservation. Uncovering of these sites requires testing a large collection of geographically diverse strains to differentiate between strain- and species-specific variation.

CONCLUSION

Mitogenomes show promise for identification purposes of important cryptic species like *F.g.* in the FGSC. Improved identification could be achieved by the combination of intronic variation analysis and whole-mitogenome SNP analysis. However, mitogenomes also show evidence of ancestral gene flow among other members of the FGSC, which largely limits clustering-based identification.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

JW and TK wrote the manuscript. JW, AD, MŻ, AS, and KB contributed to isolation and strain growth. JW, MŻ, and KB extracted the fungal DNA and prepared libraries for whole-genome sequencing. TK assembled the mitogenomes and was responsible for the study design. KM and TM performed the mitogenome annotation. TM performed the phylogenetic analysis. AD and AS edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the National Science Center, Poland, grant numbers 2015/19/B/NZ9/01329 and 2018/31/B/NZ9/00546, and "Development Program of the University of Warmia and Mazury in Olsztyn," co-financed by the European Union under the European Social Fund from the Operational Program Knowledge Education Development. JW is a recipient of a scholarship from the program "Interdisciplinary Doctoral Studies in Biology and Biotechnology" (project number POWR.03.05.00-00-Z310/17), which is funded by the European Social Fund.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.714651/full#supplementary-material

REFERENCES

Al-Reedy, R. M., Malireddy, R., Dillman, C. B., and Kennell, J. C. (2012). Comparative analysis of *Fusarium* mitochondrial genomes reveals a highly variable region that encodes an exceptionally large open reading frame. *Fungal Genet. Biol.* 49, 2–14. doi: 10.1016/j.fgb.2011.11.008

Bae, S. H., Oh, J. H., and Lee, J. (2021). Identification of interspecific and intraspecific single nucleotide polymorphisms in *Papaver* spp. *Plant Breed Biotechnol.* 9, 55–64. doi: 10.9787/PBB.2021.9.1.55

Basse, C. W. (2010). Mitochondrial inheritance in fungi. *Curr. Opin. Microbiol.* 13, 712–719. doi: 10.1016/j.mib.2010.09.003

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Brankovics, B., van Dam, P., Rep, M., de Hoog, G. S., van der Lee, T. A. J., Waalwijk, C., et al. (2017). Mitochondrial genomes reveal recombination in the presumed asexual *Fusarium oxysporum* species complex. *BMC Genomics* 18:735. doi: 10.1186/s12864-017-4116-5

Brankovics, B., van Diepeningen, A. D., de Hoog, G. S., van der Lee, T. A. J., and Waalwijk, C. (2020). Detecting introgression between members of the Fusarium fujikuroi and F. oxysporum species complexes by comparative mitogenomics. *Front. Microbiol.* 11:1092. doi: 10.3389/fmicb.2020.01092

Darling, A. E., Mau, B., and Perna, N. T. (2010). Progressivemauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. doi: 10.1371/journal.pone.0011147

Dean, R., Van Kan, J. A. L., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., et al. (2012). The top 10 fungal pathogens in molecular plant pathology. *Mol. Plant Pathol.* 13, 414–430. doi: 10.1111/j.1364-3703.2011.00783.x

Dierckxsens, N., Mardulyn, P., and Smits, G. (2017). NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45:18. doi: 10.1093/nar/gkw955

Dong, F., Xu, J., Zhang, X., Wang, S., Xing, Y., Mokoena, M. P., et al. (2020). Gramineous weeds near paddy fields are alternative hosts for the *Fusarium graminearum* species complex that causes fusarium head blight in rice. *Plant Pathol.* 69, 433–441. doi: 10.1111/ppa.13143

Fourie, G., van der Merwe, N. A., Wingfield, B. D., Bogale, M., Tudzynski, B., Wingfield, M. J., et al. (2013). Evidence for inter-specific recombination among the mitochondrial genomes of *Fusarium* species in the *Gibberella fujikuroi* complex. *BMC Genomics* 14:605. doi: 10.1186/1471-2164-14-605

Fourie, G., van der Merwe, N. A., Wingfield, B. D., Bogale, M., Wingfield, M. J., and Steenkamp, E. T. (2018). Mitochondrial introgression and interspecies recombination in the *Fusarium fujikuroi* species complex. *IMA Fungus* 9, 37–63. doi: 10.5598/imafungus.2018.09.01.04

Franco, M. E. E., López, S. M. Y., Medina, R., Lucentini, C. G., Troncozo, M. I., Pastorino, G. N., et al. (2017). The mitochondrial genome of the plant-pathogenic fungus *Stemphylium lycopersici* uncovers a dynamic structure due to repetitive and mobile elements. *PLoS One* 12:e0185545. doi: 10.1371/journal.pone.0185545

Guha, T. K., Wai, A., Mullineux, S. T., and Hausner, G. (2018). The intron landscape of the mtDNA cytb gene among the *Ascomycota*: introns and intron-encoded open reading frames. Mitochondrial DNA part A DNA mapping. *Seq. Anal.* 29, 1015–1024. doi: 10.1080/24701394.2017.1404042

Jianbo, Q. (2014). Molecular characterization of the *Fusarium graminearum* species complex in eastern China. *Artic. Eur. J. Plant Pathol.* 139, 811–823. doi: 10.1007/s10658-014-0435-4

Joardar, V., Abrams, N. F., Hostetler, J., Paukstelis, P. J., Pakala, S., Pakala, S. B., et al. (2012). Sequencing of mitochondrial genomes of nine *Aspergillus* and *Penicillium* species identifies mobile introns and accessory genes as main sources of genome size variability. *BMC Genomics* 13:698. doi: 10.1186/1471-2164-13-698

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden, T. L. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Res.* 36, 5–9. doi: 10.1093/nar/gkn201

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/MOLBEV/MST010

Katoh, K., and Toh, H. (2010). Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics* 26, 1899–1900. doi: 10.1093/bioinformatics/btq224

Kelly, A. C., and Ward, T. J. (2018). Population genomics of *Fusarium graminearum* reveals signatures of divergent evolution within a major cereal pathogen. *PLoS One* 13:e0194616. doi: 10.1371/journal.pone.0194616

Kolesnikova, A. I., Putintseva, Y. A., Simonov, E. P., Biriukov, V. V., Oreshkova, N. V., Pavlov, I. N., et al. (2019). Mobile genetic elements explain size variation in the mitochondrial genomes of four closely-related *Armillaria* species. *BMC Genomics* 20:351. doi: 10.1186/s12864-019-5732-z

Krimitzas, A., Pyrri, I., Kouvelis, V. N., Kapsanaki-Gotsi, E., and Typas, M. A. (2013). A phylogenetic analysis of greek isolates of *Aspergillus* species based on morphology and nuclear and mitochondrial gene sequences. *Biomed. Res. Int.* 2013, 1–18. doi: 10.1155/2013/260395

Kulik, T., Bilska, K., and Żelechowski, M. (2020a). Promising perspectives for detection, identification, and quantification of plant pathogenic fungi and oomycetes through targeting mitochondrial DNA. *Int. J. Mol. Sci.* 21:2645. doi: 10.3390/ijms21072645

Kulik, T., Brankovics, B., van Diepeningen, A. D., Bilska, K., Żelechowski, M., Myszczyński, K., et al. (2020b). Diversity of mobile genetic elements in the mitogenomes of closely related *Fusarium culmorum* and *F. graminearum* sensu stricto strains and its implication for diagnostic purposes. *Front. Microbiol.* 11:1002. doi: 10.3389/fmicb.2020.01002

Kulik, T., Ostrowska, A., Buśko, M., Pasquali, M., Beyer, M., Stenglein, S., et al. (2015). Development of an FgMito assay: a highly sensitive mitochondrial based qPCR assay for quantification of *Fusarium graminearum* sensu stricto. *Int. J. Food Microbiol.* 210, 16–23. doi: 10.1016/j.ijfoodmicro.2015.06.012

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., and Calcott, B. (2016). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34, 772–773. doi: 10.1093/molbev/msw260

Leekitcharoenphon, P., Kaas, R. S., Thomsen, M. C. F., Friis, C., Rasmussen, S., and Aarestrup, F. M. (2012). snpTree--a web-server to identify and construct SNP trees from whole genome sequence data. *BMC Genomics* 13(Suppl. 7):S6. doi: 10.1186/1471-2164-13-S7-S6

Losada, L., Pakala, S. B., Fedorova, N. D., Joardar, V., Shabalina, S. A., Hostetler, J., et al. (2014). Mobile elements and mitochondrial genome expansion in the soil fungus and potato pathogen *Rhizoctonia solani* AG-3. *FEMS Microbiol. Lett.* 352, 165–173. doi: 10.1111/1574-6968.12387

Luchi, N., Ioos, R., and Santini, A. (2020). Fast and reliable molecular methods to detect fungal pathogens in woody plants. *Appl. Microbiol. Biotechnol.* 104, 2453–2468. doi: 10.1007/s00253-020-10395-4

Marchler-Bauer, A., and Bryant, S. H. (2004). CD-search: protein domain annotations on the fly. *Nucleic Acids Res.* 32, W327–W331. doi: 10.1093/nar/gkh454

Megarioti, A. H., and Kouvelis, V. N. (2020). The coevolution of fungal mitochondrial introns and their homing endonucleases (GIY-YIG and LAGLIDADG). *Genome Biol. Evol.* 12, 1337–1354. doi: 10.1093/gbe/evaa126

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015

Mitchell, A., Chang, H. Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., et al. (2015). The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* 43, D213–D221. doi: 10.1093/nar/gku1243

Pavesi, A., Conterio, F., Bolchi, A., Dieci, G., and Ottonello, S. (1994). Identification of new eukaryotic tRNA genes in genomic DNA databases by a multistep weight matrix analysis of transcriptional control regions. *Nucleic Acids Res.* 22, 1247–1256. doi: 10.1093/nar/22.7.1247

Pogoda, C. S., Keepers, K. G., Nadiadi, A. Y., Bailey, D. W., Lendemer, J. C., Tripp, E. A., et al. (2019). Genome streamlining via complete loss of introns has occurred multiple times in lichenized fungal mitochondria. *Ecol. Evol.* 9, 4245–4263. doi: 10.1002/ece3.5056

Santamaria, M., Vicario, S., Pappadà, G., Scioscia, G., Scazzocchio, C., and Saccone, C. (2009). Towards barcode markers in fungi: an intron map of Ascomycota mitochondria. *BMC Bioinform.* 10(Suppl. 6):S15. doi: 10.1186/1471-2105-10-S6-S15

Sarver, B. A. J., Ward, T. J., Gale, L. R., Broz, K., Corby Kistler, H., Aoki, T., et al. (2011). Novel Fusarium head blight pathogens from Nepal and Louisiana revealed by multilocus genealogical concordance. *Fungal Genet. Biol.* 48, 1096–1107. doi: 10.1016/j.fgb.2011.09.002

Schürch, A. C., Arredondo-Alonso, S., Willems, R. J. L., and Goering, R. V. (2018). Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene–based approaches. *Clin. Microbiol. Infect.* 24, 350–354. doi: 10.1016/j.cmi.2017.12.016

Uelze, L., Borowiak, M., Bönn, M., Brinks, E., Deneke, C., Hankeln, T., et al. (2020a). German-wide interlaboratory study compares consistency, accuracy and reproducibility of whole-genome short read sequencing. *Front. Microbiol.* 11:573972. doi: 10.3389/fmicb.2020.573972

Uelze, L., Grützke, J., Borowiak, M., Hammerl, J. A., Juraschek, K., Deneke, C., et al. (2020b). Typing methods based on whole genome sequencing data. *One Heal. Outlook* 2, 1–19. doi: 10.1186/s42522-020-0010-1

van der Lee, T., Zhang, H., van Diepeningen, A., and Waalwijk, C. (2015). Biogeography of *Fusarium graminearum* species complex and chemotypes: a review. *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess* 32, 453–460. doi: 10.1080/19440049.2014.984244

Walkowiak, S., Rowland, O., Rodrigue, N., and Subramaniam, R. (2016). Whole genome sequencing and comparative genomics of closely related Fusarium head blight fungi: *Fusarium graminearum*, *F. meridionale* and *F. asiaticum*. *BMC Genomics* 17:1014. doi: 10.1186/s12864-016-3371-1

Wolters, J. F., Chiu, K., and Fiumera, H. L. (2015). Population structure of mitochondrial genomes in *Saccharomyces cerevisiae*. *BMC Genomics* 16:451. doi: 10.1186/s12864-015-1664-4

Wu, K.-L., Chen, W.-Z., Yang, S., Wen, Y., Zheng, Y.-R., Anjago, W. M., et al. (2019). Isolation and identification of *Fusarium oxysporum f. sp. cubense* in Fujian Province, China. *J. Integr. Agric.* 18, 1905–1913. doi: 10.1016/S2095-3119(18)62149-5

Wu, M., Kostyun, J. L., Hahn, M. W., and Moyle, L. C. (2018). Dissecting the basis of novel trait evolution in a radiation with widespread phylogenetic discordance. *Mol. Ecol.* 27, 3301–3316. doi: 10.1111/mec.14780

Yang, M., Zhang, H., van der Lee, T. A. J., Waalwijk, C., van Diepeningen, A. D., Feng, J., et al. (2020). Population genomic analysis reveals a highly conserved mitochondrial genome in *Fusarium asiaticum*. *Front. Microbiol.* 11:839. doi: 10.3389/fmicb.2020.00839

Yao, L., Li, Y., Ma, C., Tong, L., Du, F., and Xu, M. (2020). Combined genome-wide association study and transcriptome analysis reveal candidate genes for resistance to *Fusarium* ear rot in maize. *J. Integr. Plant Biol.* 62, 1535–1551. doi: 10.1111/jipb.12911

Zhang, C., Dai, Y., Wang, G., Wang, C., Gao, Y., Deng, W., et al. (2020). Mitogenome of *Tolypocladium guangdongense*. *Appl. Microbiol. Biotechnol.* 104, 9295–9308. doi: 10.1007/s00253-020-10889-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

REVIEW
published: 11 October 2021
doi: 10.3389/fmicb.2021.748178

Uncovering the Role of Metabolism in Oomycete–Host Interactions Using Genome-Scale Metabolic Models

Sander Y. A. Rodenburg[1,2‡], Michael F. Seidl[1,3†], Dick de Ridder[2†] and Francine Govers[1,†]*

[1]*Laboratory of Phytopathology, Wageningen University & Research, Wageningen, Netherlands,* [2]*Bioinformatics Group, Wageningen University & Research, Wageningen, Netherlands,* [3]*Theoretical Biology & Bioinformatics group, Department of Biology, Utrecht University, Wageningen, Netherlands*

Metabolism is the set of biochemical reactions of an organism that enables it to assimilate nutrients from its environment and to generate building blocks for growth and proliferation. It forms a complex network that is intertwined with the many molecular and cellular processes that take place within cells. Systems biology aims to capture the complexity of cells, organisms, or communities by reconstructing models based on information gathered by high-throughput analyses (omics data) and prior knowledge. One type of model is a genome-scale metabolic model (GEM) that allows studying the distributions of metabolic fluxes, i.e., the "mass-flow" through the network of biochemical reactions. GEMs are nowadays widely applied and have been reconstructed for various microbial pathogens, either in a free-living state or in interaction with their hosts, with the aim to gain insight into mechanisms of pathogenicity. In this review, we first introduce the principles of systems biology and GEMs. We then describe how metabolic modeling can contribute to unraveling microbial pathogenesis and host–pathogen interactions, with a specific focus on oomycete plant pathogens and in particular *Phytophthora infestans*. Subsequently, we review achievements obtained so far and identify and discuss potential pitfalls of current models. Finally, we propose a workflow for reconstructing high-quality GEMs and elaborate on the resources needed to advance a system biology approach aimed at untangling the intimate interactions between plants and pathogens.

Keywords: genome-scale metabolic model, systems biology, metabolic networks, *Phytophthora infestans*, plant pathogenic oomycetes, plant–pathogen interactions

OPEN ACCESS

Edited by:
Danyu Shen,
Nanjing Agricultural University, China

Reviewed by:
Chuan Xu,
Shanghai Jiao Tong University, China
Babak Momeni,
Boston College, United States

***Correspondence:**
Francine Govers
francine.govers@wur.nl

[†]These authors share senior authorship

‡Present address:
Sander Y. A. Rodenburg,
The Hyve B.V., Utrecht, Netherlands

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *27 July 2021*
Accepted: *10 September 2021*
Published: *11 October 2021*

Citation:
Rodenburg SYA, Seidl MF, de Ridder D and Govers F (2021) Uncovering the Role of Metabolism in Oomycete–Host Interactions Using Genome-Scale Metabolic Models. Front. Microbiol. 12:748178. doi: 10.3389/fmicb.2021.748178

INTRODUCTION

The metabolism of an organism defines its capabilities to take up nutrients from the environment and to convert these into essential building blocks such as nucleic acids and amino acids (Lazar and Birnbaum, 2012). Cellular metabolism can be described as a system of biochemical conversions (reactions), most of which are catalyzed by enzymes. Prokaryotic and eukaryotic genomes can encode hundreds to a few thousand metabolic enzymes (Yilmaz and Walhout, 2017). Each enzyme acts on a selection of substrates and converts these into products, typically by adding or removing reactive groups. Reactions that share substrates or products can be considered functionally connected. The collection of biochemical reactions within a cell

thus forms a large, interconnected network that represents the routes by which an organism converts simple nutrients into complex metabolites and vice versa. This network is distributed over different subcellular compartments (organelles), and transporter proteins as well as channels facilitate the transport of metabolites across lipid bilayers that surround the cell and the organelles (Sahoo et al., 2014). The overall system is subject to many parameters, such as variability in substrate concentrations, temperature, or the pH, not only in the extracellular space but also within cells. Cells regulate this system to maintain homeostasis, i.e., the ability to perform important cellular functions despite variations (perturbations), which provides robustness (Eberl, 2018; Nijhout et al., 2018). The ability to sense environmental variations and metabolic cues, and to adapt metabolism accordingly, depends on a tightly interlinked regulatory system that involves feedback loops embedded in interaction networks crossing metabolic, protein, transcript, and (epi) genetic levels (**Figure 1**; Watson et al., 2015). As such, the phenotype of the cell is an emergent property of the system's complexity (Aderem, 2005). The rates of individual metabolic reactions are tightly linked to the overall state of cellular metabolism, and therefore, understanding a small part of the system (e.g., a single enzyme or pathway) provides only limited insight into the complete system. Thus, holistic approaches are essential to understand how the state of a system can lead to the complex phenotype of an organism. Systems biology is a discipline based on such holistic approaches.

In this review, we first provide a broad overview of the field of systems biology, with a focus on genome-scale metabolic models (GEMs). We then discuss the relevance of GEMs to study pathogens and host–pathogen interactions, particularly oomycete plant pathogens such as *Phytophthora infestans*. We provide an overview of recent developments in this field and discuss challenges in reconstructing GEMs in these organisms. Finally, we propose a workflow for reconstructing high-quality GEMs and lay out a number of challenges that need to be addressed for systems biology to provide its full potential to study the intimate interactions between plants and pathogens.

SYSTEMS BIOLOGY PROVIDES A HOLISTIC OVERVIEW

The rise of computational biology, high-throughput analyses tools, and advanced measurement technologies in the early twenty-first century enabled biologists to analyze the presence of basically all molecules in an organism and to measure their quantities and interactions at cell or tissue level (Reed et al., 2006). Today, the genomes of numerous species have

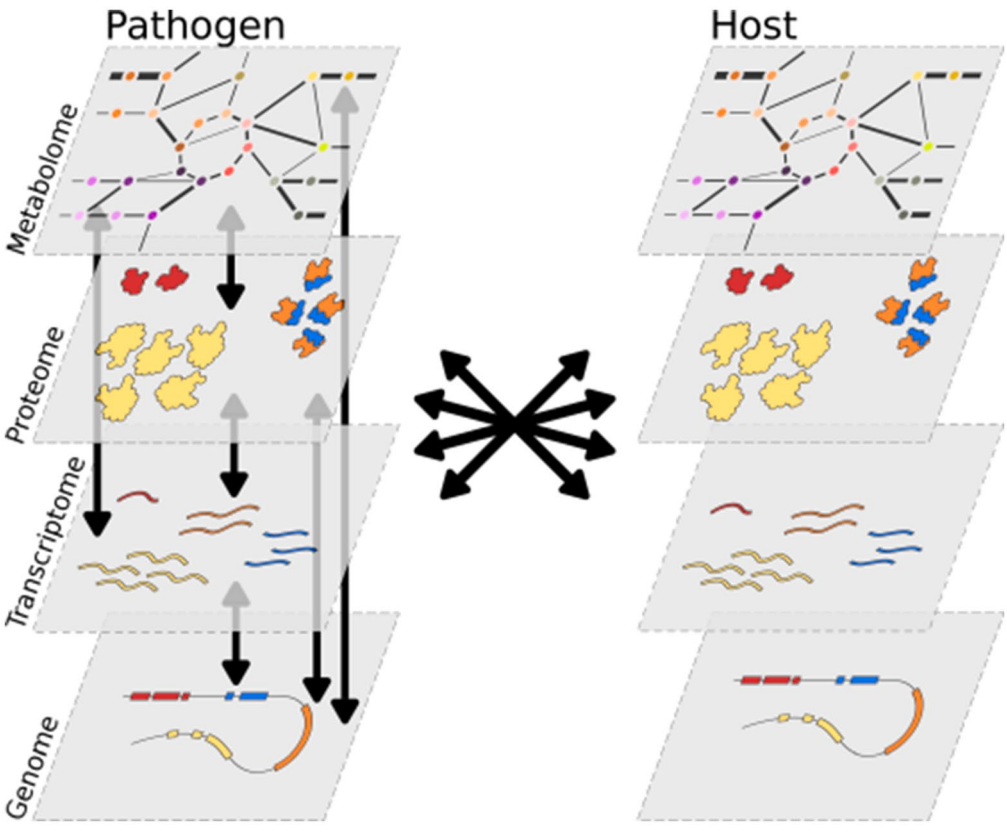


FIGURE 1 | The molecular layers of a cell are all interconnected and form a complex and integrated system. In the symbiosis between a pathogen and a host, their systems are connected, and interactions occur between all molecular layers.

been sequenced, and transcriptome sequencing as a proxy for gene expression levels has become routine. Additionally, high-throughput methods such as mass spectrometry allow the analyses of the proteome and metabolome. Data analyses typically focus on differential abundances of transcripts, proteins, or metabolites between samples. Alternatively, omics data can be used to construct models of molecular systems, as part of a scientific discipline known as systems biology (Bordbar et al., 2014) that integrates—among others—bioinformatics, mathematics, biochemistry, molecular biology, physics, and engineering (Breitling, 2010). In systems biology, omics data are integrated with prior experimental knowledge into models of molecular interactions at genome scale, aiming to capture the complete molecular systems that result in the phenotype of an organism (Yurkovich and Palsson, 2016). This includes models of protein–protein interactions, metabolic fluxes, regulatory interactions, or signaling pathways (Albert, 2007). In systems biology, high-throughput data are integrated into computational models that describe the state of the whole system.

MODELING METABOLISM

Many different types of models exist, such as ordinary differential equation (ODE)-based models for enzyme kinetics, Bayesian, Boolean, or rule-based models for signaling and regulatory networks, or constraint-based steady-state models of metabolism such as GEMs; **Box 1**; Bordbar et al., 2014; Bartocci and Lió, 2016). Metabolism is arguably the best described cellular system, and the availability of reaction information (i.e., the conversions catalyzed by metabolic enzymes) makes GEMs particularly powerful tools to investigate this system and to model metabolic fluxes within an organism (DeBerardinis and Thompson, 2012; Gu et al., 2019). They are based on the predicted enzyme repertoire found in the genome sequence, and the associated network of biochemical reactions with substrates and products. When assuming that enzymatic activity and substrate specificity of orthologs are conserved, the established metabolism in model organisms can serve as a Rosetta Stone for other organisms (Ideker et al., 2001). GEMs allow the integration of miscellaneous omics data and prior knowledge about the metabolic properties of an organism (Zhang and Hua, 2016). Furthermore, they can facilitate *in silico* identification of essential genes, reactions, or metabolites by predicting the effect of enzyme knockouts on the functioning of the system (Chavali et al., 2012). Since metabolism is profoundly connected to all other systems in the cell, a GEM can be used as a proxy to describe the phenotypic state of an organism and serves as a framework to guide future experimental research (McKnight, 2010).

Because of their integrative nature, by relating genes with reactions and metabolites, GEMs can serve as a knowledge base for species-specific information on the biochemical capacity of an organism, as deduced from its genome and from prior knowledge. Prior knowledge typically integrated in a GEM includes, for instance, the nutrients taken up and the metabolites produced in the form of biomass or secondary metabolites. Importantly, prior knowledge should be used to correct the model where automated methods are limited, such as the inference of species-specific enzymatic substrates, gene–protein–reaction associations (i.e., which genes catalyze what reactions, including subunits and isozymes), subcellular localization of enzymes, directionality of reactions, and substrates of transporters (Thiele and Palsson, 2010). Next to a knowledge base, GEMs provide a scaffold for the integration of additional omics data (O'Brien et al., 2015). A cornerstone of systems biology is the continuous integration of new data and new information into existing GEMs, with the aim to improve the quality of a GEM and make it more trustworthy.

The main purpose of GEMs is enabling flux simulations that can be used to investigate system complexity and dynamics (Orth et al., 2010). For instance, this can help unravel which flux distributions are thermodynamically optimal for growth (biomass production)—sometimes in various transcriptomic

BOX 1 | Genome-scale metabolic models

Metabolism is a complex system of thousands of biochemical reactions, most of which are catalyzed by metabolic enzymes. GEMs have been developed as an effective way to generate hypotheses about cellular metabolism (Yurkovich and Palsson, 2016). A GEM is based on the repertoire of metabolic enzymes encoded in the genome, which are cross-referenced with biochemical databases or template models to obtain a set of biochemical reactions that can transform substrate metabolites into products. These reactions are interconnected by shared substrates or products, forming a complex network that is typically divided over several subcellular compartments, such as the mitochondria and the cytosol. The metabolic network can be represented by a sparse integer matrix, denoting the stoichiometry of substrate and product metabolites of each reaction.

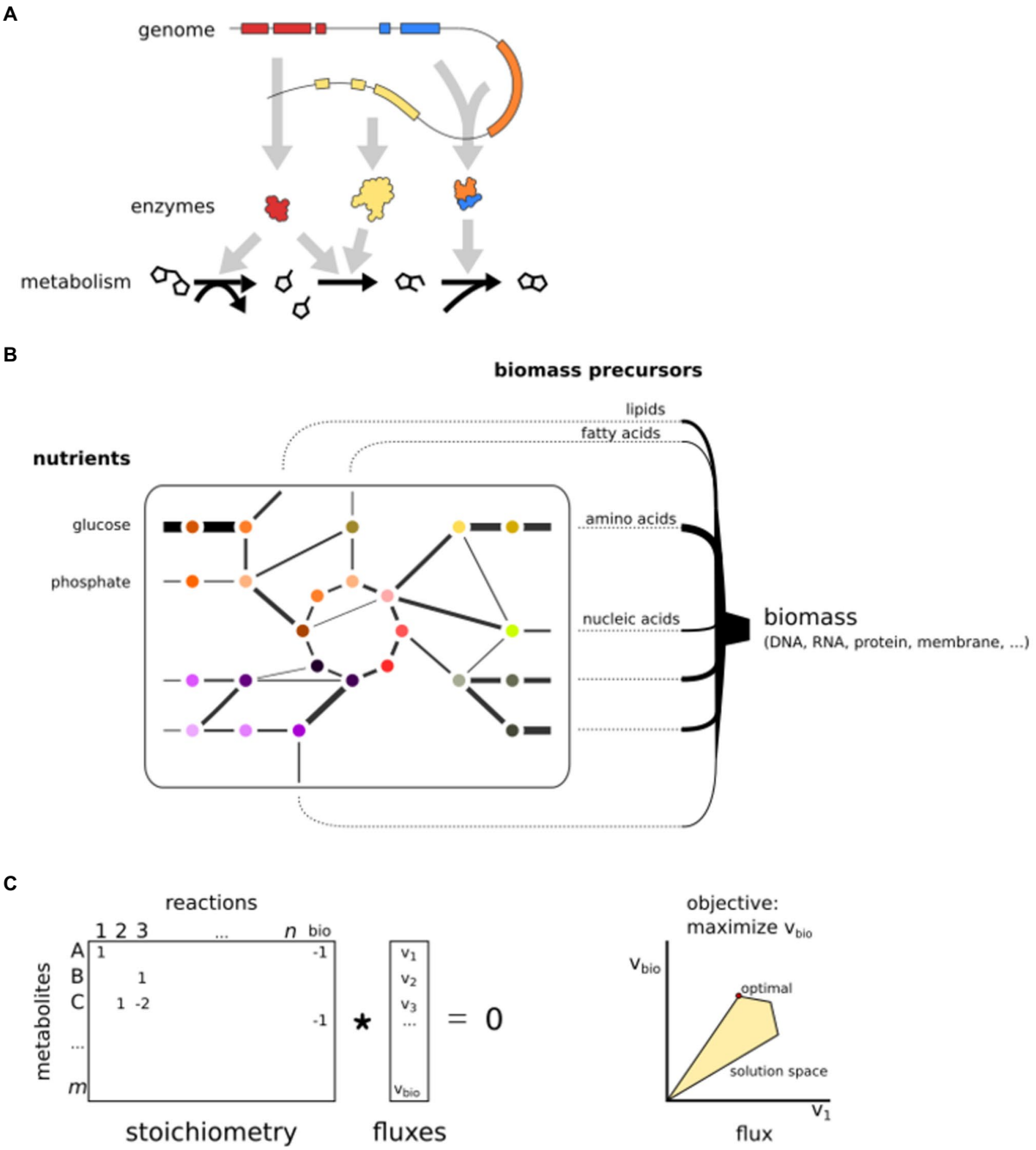
GEMs are assumed to be in steady state, which means the net uptake of nutrient mass is equal to the net production of biomass, implying there is no net accumulation of metabolites. This simplifies the model to a system of linear equations. Each reaction in the GEM is considered to have a flux, i.e., a steady-state reaction rate. The most popular method to model metabolism in this framework is called flux balance analysis (FBA; Orth et al., 2010). An assumption is that cells have a specific objective, often maximization of growth (production of biomass) or minimal usage of energy (García Sánchez and Torres Sáez, 2014). Linear optimization can find values for all fluxes that attain the specified objective, for instance, maximal flux toward biomass precursors. The production of biomass is modeled as a pseudo-reaction that consumes all biomass precursors (e.g., amino acids for proteins) with appropriate stoichiometry, often manually implemented based on experimental data (Feist and Palsson, 2010). As there is an infinite number of solutions that allow biomass production, constraints need to be implemented to find a single set of fluxes. Thermodynamic constraints can implement an upper and lower bound on each flux, specifying that reactions are either bidirectional (flux can be either negative or positive) or irreversible (Either the upper or lower bound is zero.) These constraints significantly limit the number of possible outcomes of the optimization problem. The flux values that yield a maximal value for the specified objective, within the constraints, are then selected as an optimal solution.

The integration of omics data can be used to impose additional model constraints or objective functions (Bordbar et al., 2014). Since most reactions of a GEM are associated with one or more genes, optimization can take into account gene expression and calculate the fluxes that concur with underlying gene expression. For instance, the INIT algorithm (Agren et al., 2012) maximizes a global score, which increases when "expressed" reactions have a flux and

(Continued)

decreases when "non-expressed" reactions have a flux. As a result, the optimal solution depends on the expression of genes (i.e., context-specific submodels), which enables the comparison of fluxes between transcriptome conditions. Similarly, metabolomics data can be used to calculate fluxes that attain the presence of measured metabolites. Note that these are only two specific examples of omics integration into GEMs, in a rapidly expanding set of methods (Lewis et al., 2012).

The linear program of a GEM can often be solved in a matter of seconds on modern computers, making it a powerful computational tool to predict the effect of perturbations to GEMs (Peyraud et al., 2018). For instance, the impact of a different growth medium can be analyzed (i.e., change of nutrient uptake reactions), or genes/reactions can be iteratively removed from the model to investigate the effect on the metabolic fluxes. Reactions that have large effects on the flux distribution or biomass production upon removal suggest biological relevance (Chavali et al., 2012). Essential genes or reactions can be predicted when their elimination yields an infeasible problem, i.e., no solutions respecting the implemented constraints (e.g., no biomass flux possible after removal; Pratapa et al., 2015). Similarly, synthetic lethal gene or reaction pairs can be identified that will only impair biomass production upon simultaneous deletion.



Basic principles of GEMs. **(A)** GEMs are reconstructed from a genome sequence and connect enzymes to reactions. **(B)** A GEM is a network that simulates fluxes from nutrient uptake to the production of biomass precursors. **(C)** The mathematical representation of a GEM. Reactions are stored in a stoichiometric matrix that is multiplied by a vector of fluxes. The solution space is limited by constraints. Here, only two fluxes are shown, but the optimization is in *n* dimensions.

contexts. Moreover, these simulations can be used to investigate the system's robustness to induced perturbations (e.g., gene deletions), yielding testable model-driven hypotheses. This process is key for the cyclic process in which model predictions are experimentally validated, driving technological advance, allowing for the integration of new data, and again enabling generation of new hypotheses (Kitano, 2002). A continuous cycle of improvements with additional data and knowledge can eventually lead to a highly predictive model to provide a deeper understanding of the molecular systems of an organism.

SYSTEMS BIOLOGY OF PATHOGENS AND HOST–PATHOGEN INTERACTIONS

Systems biology offers a powerful toolbox to study pathogens and their relation with their hosts (Horn et al., 2012; Durmuş et al., 2015; Dix et al., 2016; Peyraud et al., 2017; Cesur et al., 2018). Pathogens and hosts often interact extensively on all molecular levels, i.e., metabolic, protein, and DNA/RNA (**Figure 1**). Co-evolution shapes the host's immune system to be able to recognize the presence or action of a pathogen and to activate immune responses (Cook et al., 2015). To counter these processes and to facilitate infection, pathogens secrete virulence factors in the form of proteins (effectors), small RNAs, or (secondary) metabolites (Frantzeskakis et al., 2019). The symbiosis between pathogen and host can be regarded as a single intertwined system separated into different compartments (Olive and Sassetti, 2016). Thus, the reconstruction of a model for the overall system can help to characterize pathogen–host interactions and their dependencies at unprecedented scale and detail.

Systems biology has been used in the study of various pathogens or pathogen–host interactions to identify drug targets or key factors that allow pathogens to interact with their host (Durmuş et al., 2015). In pathogen–host interactions, protein–protein or small RNA interaction networks have been investigated using graph theory to identify pathogen effectors and their host interactors, in which network centrality or degree is considered a proxy for functional importance (Durmuş Tekir and Ülgen, 2012). GEMs can simulate the system-wide metabolic fluxes of a pathogen and help identify important genes, reactions, and metabolites, which can inspire novel control strategies (Chavali et al., 2012). Not surprisingly, the first GEM ever generated was for a microbial pathogen, i.e., the bacterium *Haemophilus influenzae* that causes disease in humans (Edwards and Palsson, 1999). Since then, GEMs have been reconstructed for many more pathogens, such as the tuberculosis bacterium *Mycobacterium tuberculosis* (Kavvas et al., 2018; Rienksma et al., 2018) and the human and animal parasites of the genera *Plasmodium* (Plata et al., 2010; Stanway et al., 2019) and *Leishmania* (Subramanian et al., 2015; Sharma et al., 2017; Chauhan and Singh, 2019). Moreover, some GEMs integrated pathogen and host, thereby providing insight into the metabolic fluxes throughout infection (Bordbar et al., 2010; Huthmacher et al., 2010; Bazzani et al., 2012).

In contrast to human pathogens, plant pathogens have thus far hardly been studied with a systems biology approach (Peyraud et al., 2017). Similar to human pathogens, plant pathogens have a major negative impact on the well-being of their hosts. Plants are crucial for generating the oxygen (O_2) we breathe, for sequestering CO_2 and maintaining the balance in the global ecosystem, and for the production of food and feed. However, plants are under constant threat of pathogens, such as fungi, oomycetes, bacteria, and viruses. In agriculture, the resulting yield losses can be substantial, reaching up to 30% (Savary et al., 2019). To combat plant diseases, a better understanding of plant–pathogen interactions is required. Thus far, there are, however, only few examples where systems biology was applied to provide insight into the molecular mechanisms underlying plant–pathogen interactions. In one study that was based on yeast-two-hybrid screenings, a protein–protein interaction network of *Arabidopsis thaliana* and pathogens of three kingdoms uncovered that effectors from different pathogens convergently target the same host proteins (Weßling et al., 2014). In a more recent study based on mass spectrometry analyses of immunoprecipitated effector–host target protein complexes in *Nicotiana benthamiana*, the deduced protein–protein interaction network revealed the cellular vesicle trafficking machinery as a major effector-targeted process (Petre et al., 2021). In other studies, GEMs have been reconstructed for the bacterial plant pathogens *Ralstonia solanacearum*, *Xanthomonas oryzae,* and *Pectobacterium parmentieri* (Peyraud et al., 2016; Zoledowska et al., 2018; Koduru et al., 2020), and for the fungus *Sclerotinia sclerotiorum* (Peyraud et al., 2019). However, despite the abundance of omics data for many plant pathogens, very few have been analyzed from a systems biology perspective.

OOMYCETE PATHOGENS

Oomycetes are filamentous eukaryotes that resemble fungi in terms of morphology but evolved independently from fungi (McGowan and Fitzpatrick, 2020). In the tree of life, oomycetes are clustered with the brown algae and diatoms in the Stramenopile lineage (Beakes et al., 2011; Keeling and Burki, 2019). Many oomycetes are plant pathogens, while others are animal pathogens, parasitize on other microbes, or live as saprophytes (Derevnina et al., 2016b). The plant pathogenic oomycetes vary in lifestyle, including necrotrophs that swiftly kill their hosts and feed off dead plant material (Fawke et al., 2015) and biotrophs that need living host tissue to infect, feed, and proliferate. Most of the biotrophic oomycetes, such as the white rusts and downy mildews, are obligate pathogens implying that they cannot grow outside a living host (Baxter et al., 2010). They usually specialize on just one plant species and hence have a very narrow host range. Others are known as hemibiotrophs; they live as biotrophs during the initial phase of the disease cycle but switch to a necrotrophic lifestyle later on. *Phytophthora* species are mostly hemibiotrophs. To date, over 150 *Phytophthora* species have been described, all with their own specific host range, sometimes limited to one or few

plant species within one family but more often multiple plant species from different families. Well-known narrow host range species are *Phytophthora infestans* that causes late blight disease on potato and tomato, and *Phytophthora sojae,* the soybean root and stem rot pathogen. Examples of broad host range species are *Phytophthora ramorum,* the sudden oak death pathogen, *Phytophthora capsici* that causes stem and fruit rot on many vegetables, and *Phytophthora palmivora*, a devastating pathogen on tropical crops such as cacao and date palm (Kamoun et al., 2014; Perrine-Walker, 2020). The type species of the genus is *P. infestans* which caused the Irish potato famine in the mid-nineteenth century. It was initially named *Botrytis infestans* but later renamed by de Bary (1876) who provided the formal proof that this organism is the causal agent of potato late blight and coined the term *Phytophthora,* Greek for plant (phyton) and destruction (phthora). Its arrival in Europe marked the birth of plant pathology as a discipline and ever since *P. infestans* has been a favorite subject of investigation (Turner, 2005). Technological advancements in molecular genetics and genomics over the last four decades further boosted the interest in studying this important plant pathogen (Turner, 2008). Although still considered a model organism for oomycetes, in research it is losing ground to other species that have smaller genome sizes and are more amenable to genetic modification.

THE INTERPLAY BETWEEN HOST AND OOMYCETE PATHOGENS

Biotrophic plant pathogens typically adhere to leaves or roots, break physical barriers, and scavenge nutrients from their host, while suppressing the host's immune system (McDowell, 2011). They achieve this by depositing a large variety of enzymes and effector proteins in the apoplast or inside the plant cell that help in paving the way for a successful infection (Whisson et al., 2016). During infection, these biotrophs grow as filamentous hyphae inside their hosts. They colonize the apoplastic space in the leaf mesophyll and form feeding structures, so-called haustoria, inside host cells. At the site of the haustoria and the apoplastic hyphae, pathogen and host form a close interface through which effectors, enzymes, and small molecules can be exchanged. This interplay often involves a prolonged symbiosis in which the pathogen feeds off the plant for growth and reproduction (Judelson and Ah-Fong, 2018).

The ability of oomycetes to live in close symbiosis with a host drives continuous adaptations of both pathogen and host. Oomycetes have dynamic genomes that allow swift adaptation (Leesutthiphonchai et al., 2018). These genomes typically harbor hundreds of effector genes (McGowan and Fitzpatrick, 2017) Comparative genomics has revealed that obligate biotrophic pathogens have suffered extensive gene loss as a result of their biotrophic lifestyle (Kemen and Jones, 2012; Fletcher et al., 2018). This adaptive capacity facilitates the evolutionary "arms race" between oomycete effectors and host resistance genes (Wang et al., 2019b), but also allows adaptation of the core cellular machinery of pathogens, including metabolism and signal transduction, leading to various unique properties (Judelson, 2017). For instance, oomycetes have several genes encoding unique proteins with novel domain combinations (Seidl et al., 2010; van den Hoogen and Govers, 2018a), as well as a number of horizontally transferred genes coding for proteins with functions in metabolism (Richards et al., 2011). Oomycetes are osmotrophs, which means they secrete enzymes to digest large molecules (polymers) extracellularly and import the resulting small molecules as nutrients (Richards and Talbot, 2013). This process is facilitated by a broad array of transporter proteins, suggesting that a plethora of host compounds can be taken up during infection (Abrahamian et al., 2016). However, some nutrients are indispensable for oomycetes. For instance, several oomycetes lost the ability to synthesize sterols (Wang et al., 2021). These sterol auxotrophs secrete elicitins, oomycete-specific proteins that are thought to be sterol carriers and likely exploited for recruiting sterols from the environment (Derevnina et al., 2016a). Moreover, most oomycetes are auxotrophic for thiamine, a vitamin that acts as a cofactor in carbohydrate catabolism (Hohl, 1991). Culturable oomycetes can be grown *in vitro* and seem to prefer amino acids as a substrate (Hodgson, 1958; Ah-Fong et al., 2017b), but can utilize a wide variety of substances. Due to the complex nature of cellular metabolism, it is currently unclear which nutrients are more important than others, and how the differential usage of nutrients might influence infection. It is also not so easy to readily gain such knowledge. Many oomycetes are hard to culture and require complex media for *in vitro* growth, often prepared from seeds such as rye kernels, peas, or lima beans. This obviously complicates biochemical assays to investigate their metabolism, for which knowledge of the precise growth substrates is mandatory. For obligate biotrophs that exclusively grow inside their living host, it is even more challenging; unravelling the precise composition of their diets is extremely difficult if not impossible (McDowell, 2011).

Validation of the role of enzymes or transporters in metabolism by targeted mutagenesis is another challenge when investigating oomycetes. For several *Phytophthora* and *Pythium* species, successful DNA transformation has been described but transformation efficiencies are often relatively low. Until recently, functional gene analyses relied on gene silencing or overexpression of the target gene with the disadvantage that the variability in silencing or overexpression levels and potential off-target effects make phenotypic characterization of the transformants complicated and labor-intensive. A major breakthrough was the achievement by Fang and Tyler (2015) who published the first successful application of CRISPR-Cas9-mediated gene editing in an oomycete, namely *P. sojae,* and by now, this is a standard method to create gene knockouts in several *Phytophthora* species (Wang et al., 2019a; Pettongkhao et al., 2020). In *P. infestans,* however, CRISPR-Cas9-mediated gene editing was not successful (van den Hoogen and Govers, 2018b). A recent

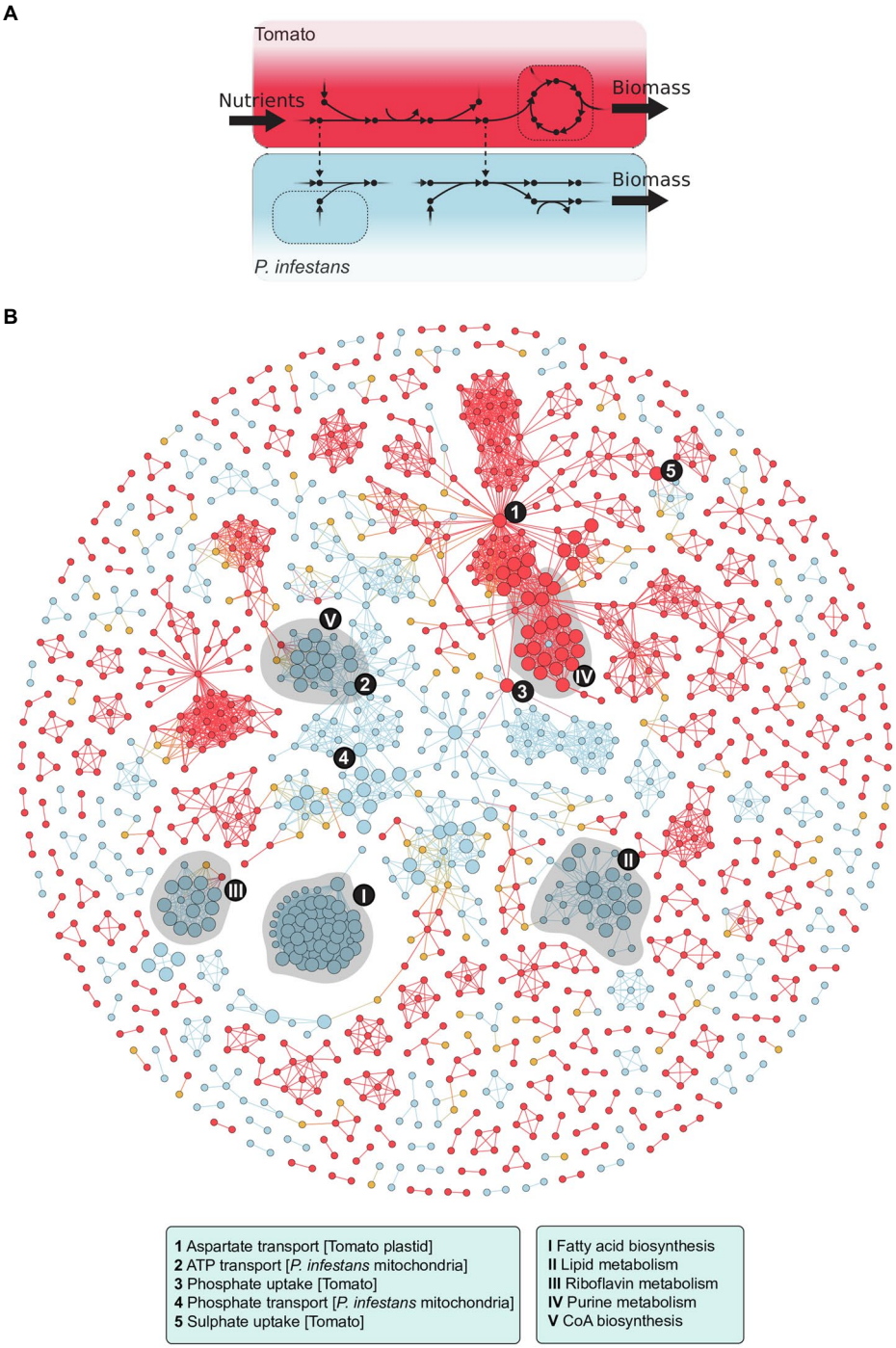


FIGURE 2 | The integrated *P. infestans*–tomato model (reproduced and adapted from Rodenburg et al., 2019). **(A)** Diagram in which dots represent metabolites, arrows reactions, and dotted lines host–pathogen transport reactions. **(B)** Flux coupling analyses of the model identified 77 coupled (i.e., topographically related) *P. infestans*–tomato reaction pairs which are shown in the graph. Nodes represent reactions in tomato (red) or *P. infestans* (blue) and host–pathogen transport (yellow). Edges represent coupling between those reactions. The nodes with the largest diameter represent 112 reactions in *P. infestans* and 35 in tomato that were found to be essential for *P. infestans* biomass production. The various processes listed in the boxes are represented by highly connected nodes (1 to 5) and in shaded clusters (I to V). Further details in Rodenburg et al. (2019).

study by Ah-Fong et al. (2021) showed that Cas9 is toxic for this species, but with Cas12a as nuclease they obtained transformants that are viable and have small deletions in the target gene, the *inf1* elicitin gene. This is another major leap forward as it demonstrates the successful implementation of a promising gene editing tool in *P. infestans*.

SYSTEMS BIOLOGY ON OOMYCETES AND OOMYCETE–HOST INTERACTIONS

Oomycete pathogens are very challenging to control due to their high capacity for adaptation. It is therefore important to reconstruct holistic models that provide mechanistic insight into the molecular systems that allow oomycetes to proliferate and infect their hosts. Ultimately, a system-wide understanding of oomycete–host interactions might provide novel leads for control (Dunphy and Papin, 2018).

Shortly after the first genome sequences of oomycetes were published (Tyler et al., 2006; Haas et al., 2009), it was already proposed to reconstruct predictive models with the aim to reveal mechanisms of oomycete–host interactions (Pinzón et al., 2009; 2010). One of the first studies that reconstructed a partial metabolic network for *Phytophthora* used the network to provide context for predicting horizontal and endosymbiotic gene transfer (Whitaker et al., 2009). In subsequent studies, Seidl et al. (2013) predicted a functional association network in *P. infestans* by projecting genomic, transcriptomic, and comparative genomic data on protein–protein interaction networks of model organisms, while Mukhtar et al. (2011), Weßling et al. (2014), and Petre et al. (2021) created holistic networks uncovering a plethora of interactions between oomycete effectors and potential host targets.

Investigation of a static metabolic network can already provide new biological insights that do not come to the foreground when only smaller subsets of the data are considered. An example is a recent study in which we identified and compared the metabolic enzymes of a broad range of oomycetes with different lifestyles and host preference and investigated their metabolic networks from an evolutionary perspective (Rodenburg et al., 2020). Similar to Thines et al. (2020), we observed lineage-specific pathway loss, and convergent loss of metabolic enzymes in obligate biotrophs reflecting their reduced metabolic capacity and greater host dependency. Intriguingly, the gene losses predominantly affected the periphery of the metabolic network, an insight that remains hidden when solely comparing genomes.

MODELING *PHYTOPHTHORA INFESTANS* METABOLISM TO PREDICT PATHOGEN–HOST INTERACTIONS

In 2018, we (Rodenburg et al., 2017) and Botero et al. (2018a) presented the first GEMs of P. infestans. Despite the slightly different reconstruction and analysis approaches and the more extensive transcriptome data set in Botero et al. (2018a) including *in planta* life stages, the overall findings were comparable. For instance, the models both pinpointed fatty acid biosynthesis as a key process in oomycetes and in both models, condition-specific metabolic patterns were apparent.

As input for our modeling (Rodenburg et al., 2017), we extracted information on *P. infestans* metabolism from the literature and identified all putative enzymes encoded in its genome by homology-based enzyme annotation. We then divided reactions over the subcellular compartments and inspected the model topology to gain insight into the biochemical processes in each compartment. A further refinement was the integration of transcriptome data; the resulting life stage-specific models showed a sharp contrast in metabolic activity between sporangia and hyphae. In oomycetes, the sporangia, which are asexual spores, likely rely on stored nutrient reserves, such as glucans and fatty acids, that are catabolized for energy production (Judelson, 2017). When sporangia disperse and reach a suitable plant surface, zoospores are released and encyst. The cysts then germinate and form an appressorium-like structure at the tip of the germ tube.

For reconstruction of a GEM, information on growth phenotypes on defined substrates is pivotal and the availability of knockout mutants would greatly contribute to validate the predictions (Nakahigashi et al., 2009; Monk et al., 2014). Unfortunately, experimental data on metabolism in *P. infestans* are very limited and, as described above, tools for knockout mutagenesis still have to be further optimized. The information that is available in the literature includes data on minimal *in vitro* growth substrates (Hohl, 1991), verified subcellular localizations of enzymes (López-Calcagno et al., 2009; Abrahamian et al., 2017), and capacity to produce a mixture of long-chain polyunsaturated fatty acids (Griffiths et al., 2003; Sun et al., 2012). Perhaps the most important limitation that we faced when reconstructing the *P. infestans* GEM was the lack of knowledge on biomass composition, i.e., the stoichiometry of *P. infestans* biomass precursors (Feist and Palsson, 2010). The biomass composition relates the fluxes in the model to a hypothetical growth rate, and as such, it can be used as a proxy for metabolic fitness. Because a precise description of *P. infestans* biomass composition was not available, we estimated it from the literature but ignored relative abundance (stoichiometry; Rodenburg et al., 2017). This rendered quantitative flux predictions infeasible, but still allowed us to investigate the model for connectivity and importance of different nutrients (Rodenburg et al., 2019). Similar challenges were faced by others modeling pathogens. Tymoshenko et al. (2015), who published a GEM for the human parasite *Toxoplasma gondii*, also reconstructed a biomass composition from the literature, ignoring stoichiometry. For a GEM of *Leishmania donovani*, Sharma et al. (2017) chose to infer the biomass composition from a *Plasmodium* GEM. For oomycetes, it may be an option to adopt the biomass composition from curated GEMs of closely related organisms, such as the brown algae *Phaeodactylum tricornutum* (Levering et al., 2016) or *Ectocarpus siliculosus* (Prigent et al., 2014), but this should be weighted against the risk of introducing new biases and uncertainties. After all, the similarity of biomass composition between brown algae and oomycetes is unknown and, further complicating matters, the biomass composition of *P. infestans* in different life stages appears to be radically different (Grenville-Briggs et al., 2008).

Although the GEM of a plant pathogen is in principle suitable to predict essential metabolic genes and reactions, and to simulate growth (biomass production) *in vitro*, it is less

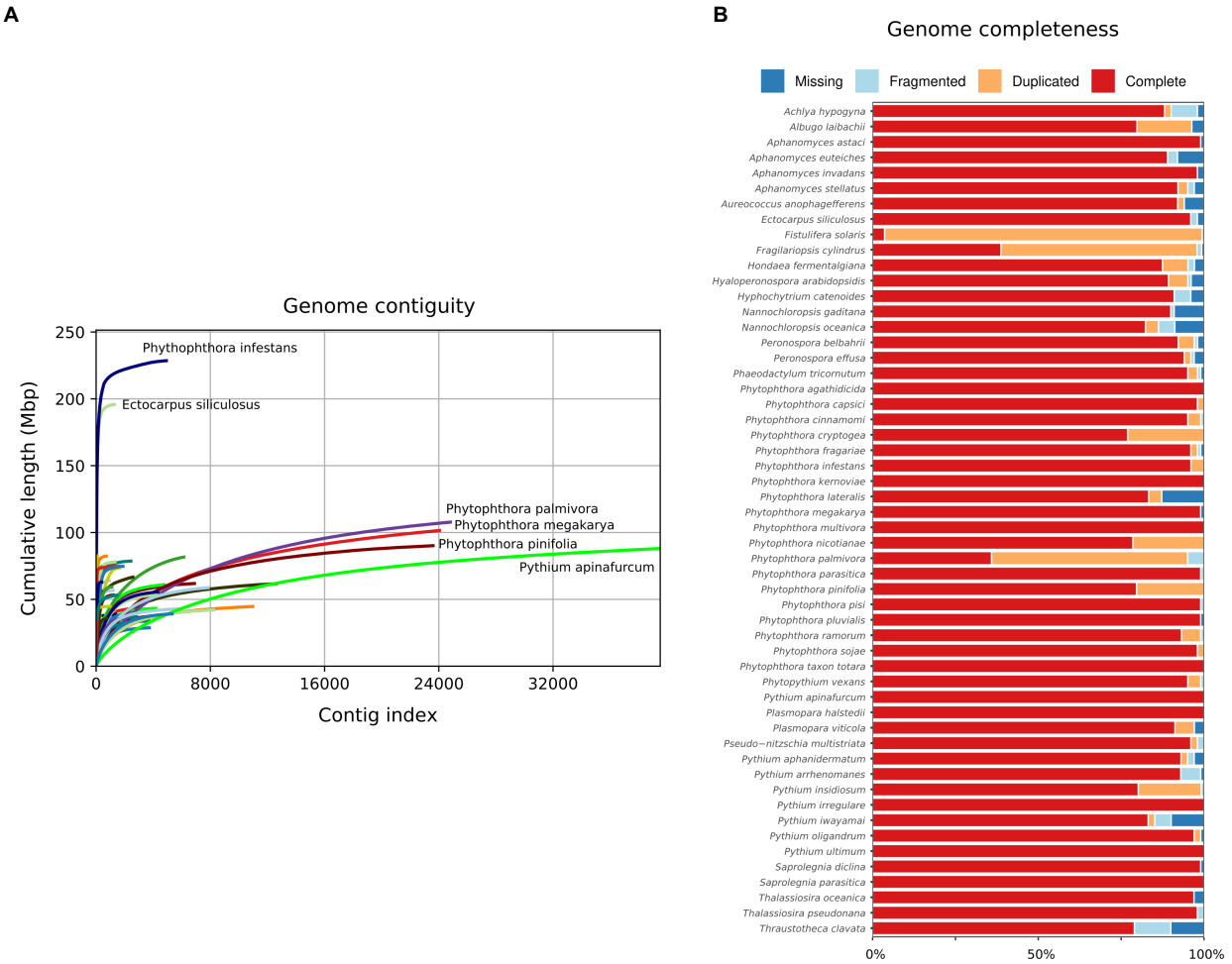


FIGURE 3 | Statistics for 54 published Stramenopile genome sequences including 42 oomycetes. **(A)** Genome accumulation curves derived from Quast (Gurevich et al., 2013), showing cumulative genome size when contigs are ordered from large to small. X-axis represents the number of contigs, and Y-axis represents genome size. **(B)** Presence of near-universal single copy orthologs in the genomes, as determined by BUSCO. Species are in alphabetical order. Colors indicate the completeness of the detected BUSCO genes. For further details on the genome sequences used for these analyses see Rodenburg et al. (2020).

informative for predicting the pathogen's metabolism during *in planta* growth. We addressed this by integrating our initial *P. infestans* GEM (Rodenburg et al., 2017) with a tomato GEM published by Yuan et al. (2016), resulting in a multi-compartment metabolic model of the *P. infestans*–tomato interaction (**Figure 2A**; Rodenburg et al., 2019). We used modeling techniques such as flux coupling analysis, to identify metabolic reactions in tomato that are of importance to fluxes in *P. infestans* metabolism. An example is thiamine biosynthesis that is required to supply the thiamine auxotrophic pathogen with this essential vitamin (Rodenburg et al., 2017). In this way, we build a GEM of a pathosystem by approaching it as a single system and demonstrated that this GEM can be used to predict metabolic changes in both host and pathogen (**Figure 2B**). Botero et al. (2018b) used an alternative approach: They constructed a GEM for potato (*Solanum tuberosum*) and modeled the metabolic changes in the plant when challenged by *P. infestans,* based on transcriptome data.

Nutrient uptake by a pathogen depends on location and stage of infection; however, a GEM of a pathosystem typically lacks resolution to take these factors into account. The first close encounter between a *Phytophthora* pathogen and a plant is when a germ tube emerging from a sporangium or cyst senses a suitable surface and starts swelling at the tip. Host entry by *Phytophthora* is facilitated by a mechanical slicing mechanism that breaches the epidermal cells (Bronkhorst et al., 2021). Haustoria emerge from hyphae that colonize the apoplast and enter the mesophyll cells (Judelson, 2017). It is often assumed that these haustoria are the main site of nutrient uptake, as is the case for various plant pathogenic fungi (Wang et al., 2018). However, many oomycetes do not form haustoria (Fawke et al., 2015), and haustoria make up only a very small proportion of the total hyphal biomass (~2%), raising the question whether haustoria are truly the main site of nutrient uptake (Judelson and Ah-Fong, 2018). The plant apoplast is a nutrient-rich environment and might be the main site of

nutrient uptake (Chen, 2013). Haustoria are nonetheless very important for the host–pathogen interaction. They form the site from where the pathogen deposits so-called cytoplasmic effectors into the plant cell for suppression of immune responses (Boevink et al., 2020). The host recognizes the intracellular host–pathogen interface created by the haustorium as the site where defense responses have to be activated and, for example, relocate the nucleus to the interface (Wang et al., 2017). Next to cytoplasmic effectors, the pathogen secretes apoplastic effectors, so also in the apoplast host and pathogen interact and likely this involves exchange of signals and compounds.

Because of the specialized tasks of the haustorium and hyphae in the apoplast, hyphal cells likely have a "division of labor," which implies that the biological processes are tailored for the specific region of infection. This phenomenon was recently modeled for the fungal plant pathogen *Sclerotium sclerotiorum*, by mapping the transcriptome of the apex and the center of infection to a multi-cell GEM (Peyraud et al., 2019). In the reconstruction of the integrated *P. infestans*–tomato GEM, we did not explicitly discriminate between the different sites of infection but rather focused on changes over time (Rodenburg et al., 2019). We integrated dual-transcriptome data obtained from infected tomato leaves. A time course of a full infection cycle was sampled with intervals of 4 h during 4 days (2–6 days after inoculation), resulting in 25 submodels. This revealed various switches in metabolism and differential nutrient usage over time, with a "division of labor" of the two partners. As infection progresses, *P. infestans* performs less *de novo* synthesis of metabolites and scavenges more metabolites from tomato. This example nicely demonstrates how one can analyze transcriptome data in a system-wide context. In concordance with related transcriptome studies (Abrahamian et al., 2016; Ah-Fong et al., 2017a), the transcriptome-based submodels reflected reduced metabolic activity in the sporangial stages of *P. infestans,* and nutritional changes in the transition from a biotrophic to a necrotrophic stage of infection on tomato leaves (Rodenburg et al., 2017, 2019). Importantly, because these transcriptomic changes were analyzed in the context of a GEM, results were subject to the imposed model constraints (steady-state, reaction thermodynamics) and thereby to the topology of the metabolic network (Hyduke et al., 2013). The transcriptomic changes are interpreted in terms of ensuing differences of metabolic fluxes, and as such, this system-wide approach can be more informative than the differential expression analyses of individual genes. The integrated metabolic model provides a framework to simulate the metabolic fluxes occurring during infection and as a result, new insights in the kind of nutrients that *P. infestans* extracts from its host during the subsequent phases of the infection cycle.

Metabolic enzymes are located in various organelles, and hence, specific metabolic processes take place in different parts of the cell. In the reconstructed *P. infestans* models, this compartmentalization was taken into account to represent the spatial distribution of metabolic pathways in different subcellular compartments (Rodenburg et al., 2017, 2019). The transporters and channels responsible for transfer of metabolic substrates across membranes were modeled by integrating transport reactions (Thiele and Palsson, 2010). Transporters typically have a wide substrate range. Because of the difficulty of predicting the substrate based on protein sequence, transporters are often manually added to a GEM based on prior knowledge. This is particularly challenging when creating an integrated pathogen–host metabolic model, considering that metabolite transport across membranes is pivotal to pathogen nutrition. In modeling the *P. infestans*–tomato interaction, we chose to not manually add transport reactions (Rodenburg et al., 2019). Because too little is known about *P. infestans* nutrition *in planta*, manually adding host–pathogen transport reactions would bias fluxes toward a predefined set of nutrient transporters. Since one of our goals was to predict the nutrient pool of *P. infestans* during tomato infection, we chose to draw conclusions based on the optimal fluxes in the model. In other words, the transport reactions in our models were largely based on network topology. Depending on the objective function, the most optimal set of transporters had a nonzero flux. The downside of this approach was that we could not consider bidirectional transport, as this would imply unrestricted metabolite exchange between host and pathogen. This could lead to scenarios where the host would utilize certain *P. infestans* metabolites for profit, which, from a biological point of view, is not plausible. In reality, however, metabolite exchange is likely a two-way process, with the host providing nutrients and the pathogen secreting metabolites, for example, as waste products or virulence factors. There is a clear knowledge gap on the metabolic exchanges that *P. infestans* maintains with its environment. To fill this gap, broad substrate screening could be performed using various growth media and different time points of mycelial growth, combined with comparative metabolomics using mass spectrometry, guided by GEM predictions.

HIGH-QUALITY GENOME DATA ARE ESSENTIAL FOR RECONSTRUCTING RELIABLE GENOME-SCALE METABOLIC MODELS

GEMs are reconstructions of cellular systems that can be used to investigate the genotype–phenotype relationship: How do the genes encoded by the genome result in the complex biological system that we observe (Yurkovich and Palsson, 2016)? Quality and insights derived from genome-scale models therefore critically depend on the quality of the genome sequence and gene annotation.

Obtaining a high-quality genome assembly is still challenging, in particular for the more complex eukaryotic genomes that often have a high abundance of repetitive elements and are typically diploid, or sometimes even polyploid or aneuploid (Nagarajan and Pop, 2013). Sequencing and comparative analyses of the first oomycete genomes in 2006 (*P. sojae* and *P. ramorum*) and 2009 (*P. infestans*) revealed that these species profoundly differ in genome size and content (Tyler et al., 2006; Haas et al., 2009). *P. infestans* has a large genome compared to its close relatives, primarily due to the high abundance of transposable elements, constituting roughly 74% of its genome. Oomycete genome assemblies are often still rather fragmented (**Figure 3A**; McGowan et al., 2018),

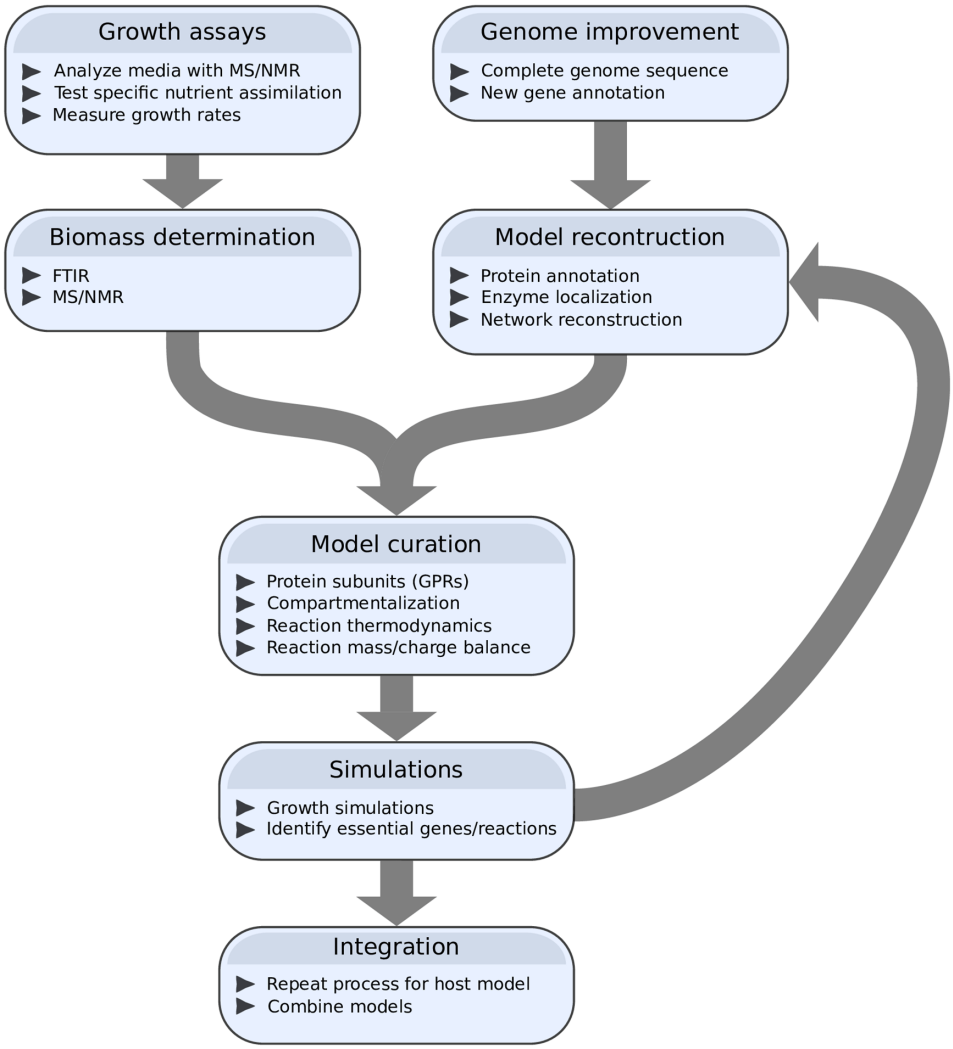


FIGURE 4 | A proposed workflow (including possible methods and analyses) for reconstructing a high-quality genome-scale metabolic model (GEM) for *P. infestans* and, potentially, an integrated GEM for a *P. infestans*/host interaction. Model reconstruction is an iterative process of simulation and model improvement.

in particular compared to fungal plant pathogens for which genome assemblies are nowadays often near-complete (Faino et al., 2015). Only over the last few years, near-complete genome assemblies have been published for some oomycetes (Fletcher et al., 2019; Malar et al., 2019; Stajich et al., 2021). Despite the fragmented genome assemblies, almost all oomycete genome sequences comprise over 95% of the near-universal single-copy conserved orthologs, as determined by BUSCO (**Figure 3B**; Seppey et al., 2019), suggesting that a significant proportion of the coding genome is captured.

Another major challenge is to correctly identify the open reading frames with associated exon boundaries (gene models) within the assembled genome sequences (Salzberg, 2019), which is even more problematic in fragmented genome assemblies (Denton et al., 2014). Eukaryotic genomes are typically annotated using gene predictors trained on the parameters of high-quality gene models from closely related species and aligned transcriptome data (Yandell and Ence, 2012). Annotation of the first two sequenced *Phytophthora* genomes was performed using a gene predictor trained on expressed sequence tags (Tyler et al., 2006). Many subsequent oomycete genome annotations were performed by gene predictors trained on the gene models in other oomycete genomes (McGowan and Fitzpatrick, 2017). However, the error rate in predicted gene models is still high, emphasizing the need for specifically searching for potentially missing gene models when mining genome sequences of oomycetes. Consequently, the field would benefit greatly from manual, and preferably community-driven, curation efforts of genome annotations (Rödelsperger et al., 2019). This would be valuable even when applied to only a single species, preferably one with a near-complete genome assembly, which can then be used as a template to re-evaluate gene models of other oomycetes and train gene predictors.

Reliable functional annotation of the predicted proteome is an additional prerequisite for identifying the fundamental components

of genome-scale models. When performing an automated proteome annotation for 54 Stramenopile species, we found that the majority of predicted proteins could not be assigned to a KEGG orthologous group (KO), a system to cluster protein orthologs with validated functions (Mao et al., 2005), and this hampered the association of these proteins with a putative function (Rodenburg et al., 2020). Many effector genes in oomycetes (e.g., those encoding RxLR effectors) lack functional associations (McGowan and Fitzpatrick, 2017). One should consider that orthologs in distantly related species are inherently harder to detect than in closely related ones, which poses a problem in homology-based annotations. However, even in the best studied eukaryotic model organism, the yeast *Saccharomyces cerevisiae,* about 20% of all genes lack any functional association (Wood et al., 2019). Lineages that are evolutionarily distant to model organisms—such as the oomycetes—have an even less functionally characterized proteome, partly because of limitations of homology-based inference of protein functions, but most importantly, because of lack of experimental characterization. An effective and sensitive tool to predict protein functions by homology is offered by hidden Markov models (HMMs). HMMs are trained on a multiple sequence alignment of a predefined cluster of homologous protein sequences and weigh conserved sequence regions heavier than variable regions. Therefore, HMMs are particularly suitable for detecting protein domains, as these are often highly conserved to retain their biological function (Pearson, 2013). KOs are predefined ortholog clusters and powerful resources to train HMMs (Aramaki et al., 2019). In our studies, we used KO-based HMMs to identify orthologs of metabolic enzymes in oomycetes and their close relatives.

In omics-based bioinformatics studies, it is common practice to search for overrepresentation (enrichment) of functional annotations in differentially abundant molecules such as mRNA or proteins (Reed et al., 2006; Bordbar et al., 2014). Over the last decade, there has been a continuous flow of studies presenting oomycete comparative genomics, transcriptomics, proteomics, proteogenomics, and metabolomics data (Andronis et al., 2020; McGowan and Fitzpatrick, 2020). These large-scale omics datasets are analyzed to understand how oomycetes evolve, reproduce, and interact with their hosts. Despite omics studies being indispensable to investigate the transcriptional and/or translational responses of both pathogen and host during infection, these studies are often biased, as usually only a subset of functions is investigated and these are not necessarily representative or causative for the complex phenotype. Moreover, proteomic and metabolomic samples typically only capture the most ubiquitous molecules, and as the differential abundance of any molecule may be influenced by subtle changes in the environmental or experimental conditions, the biological implications remain speculative.

TOWARD HIGH-QUALITY GEMS OF PATHOGENS AND PATHOGEN–HOST INTERACTIONS

Systems biology has been recognized years ago as a promising method to study plant pathogens (Pinzón et al., 2009; Pritchard and Birch, 2011). In the last few years, the potato and tomato late blight (*P. infestans*) pathosystem was subject of several systems biology studies (Seidl et al., 2013; Rodenburg et al., 2017, 2019; Botero et al., 2018a,b; Castro et al., 2019; Thines et al., 2020). Nevertheless, systems biology of this pathosystem and many others is still in its infancy. The level of knowledge on the organism to be modeled is key for the success of a systems biology approach. Thus, to arrive at highly predictive models for *P. infestans* or any other pathogen in the near future, *in vitro* experiments need to be performed to gain basal knowledge. In the case of *P. infestans*, valuable information would be, for instance, the substrates that *P. infestans* can assimilate from its environment, as well as its biomass composition and how this changes throughout its lifecycle. There is a lot we can learn from the more advanced metabolic research in other pathosystems. For example, there are now several GEMs for *Plasmodium* spp., some of which are also integrated with GEMs of the host, the red blood cell (Huthmacher et al., 2010; Plata et al., 2010; Abdel-Haleem et al., 2018). The foundation of these GEMs was provided by pathogen–host metabolomics analyses identifying growth substrates (Olszewski et al., 2009). These data were integrated with new omics data and novel biochemical knowledge into GEMs (Bazzani et al., 2012; Carey et al., 2017; Stanway et al., 2019). Interestingly, these models have pinpointed several essential reactions, some of which turned out to be leads for promising drug targets (O'Hara et al., 2014). For protozoan parasites, isotope-labeled growth experiments have been successful to dissect their metabolism during parasitic growth (Kloehn et al., 2016), and it can be anticipated that similar analyses will provide intriguing novel avenues to control oomycete and fungal plant pathogens.

To be able to take full advantage of systems biology, several steps to create a higher-quality GEM for *P. infestans* as well as other plant pathogens are needed in the future (Thiele and Palsson, 2010). Obviously, complete functional characterizations of the substrates and characteristics of each individual metabolic enzyme in the pathogen as well as in the host would be ideal, but this seems infeasible in the near future. Nevertheless, significant achievements could be gained from *in silico* and *in vitro* procedures, designed specifically for the purpose of building a high-quality GEM (**Figure 4**).

First (near-)complete and ideally gapless genome assemblies of pathogens as well as their hosts are required. The genomes need to be sequenced and assembled using novel technologies and advanced assembly methods to attain the complete coding information (Thomma et al., 2016). Gene prediction should be guided by RNA sequencing, homology-based evidence, and metabolics data, and predicted protein sequences should be functionally annotated. Moreover, manual gene model curation should be performed to accurately predict the enzymes and all this information should be linked to knowledge on the biochemical capacity of the pathogen (Fernandes et al., 2019). Such manual curation is time-and labor-intensive, but will ultimately lead to a complete description of the enzyme catalogue encoded in the genomes and thus to better models.

Second, the identified enzymes are then used as input for reconstructing a draft metabolic network. Functional annotation

of (subunits of) metabolic enzymes and the associated reactions and metabolites should be curated according to the literature and according to the established protocol (Thiele and Palsson, 2010), pinpointing gene–protein–reaction associations. The metabolic network needs to be manually inspected to identify potentially missing enzymes (gaps), by comparing the reconstructed pathway to reference pathways of well-characterized model organisms, and the annotations need to be revisited to account for apparently missing enzymes (i.e., wrong or erroneous gene annotation). Additionally, these models should be compartmentalized, i.e., reactions should be assigned to the correct cellular location with transport reactions to simulate the transmembrane metabolite fluxes. We and others previously considered at least the cytosol, extracellular space, and the mitochondria essential, since these are hotspots of metabolism. Intracellular transport reactions need to be manually included based on common biochemical knowledge, such as experimental data, textbooks, or literature. We found that specifically for oomycetes the experimental data on transporter substrates are very limited, and it is unlikely that many more transporters are to be characterized, given the labor-intensive process (Savory et al., 2018). Therefore, transport reactions may be inferred from GEMs of related species and by analyzing network topology.

Third, *in vitro* assays can be performed to characterize the pathogen's growth behavior under different conditions. The ideal medium is similar in composition to the pathogen's natural hosts to mimic natural growth. The medium needs to be analyzed by untargeted metabolomics (mass spectrometry and/or nuclear magnetic resonance) over multiple time points to provide insight into the presence and abundance of specific metabolites. Metabolites that strongly change in abundance during pathogen growth and between subsequent sampling stages are likely assimilated or secreted. This can be indicative of a nutrient transporter on the plasma membrane that is capable of transporting the respective metabolite. In addition, isotope-labeled metabolites can be added to the medium to test the assimilation of specific nutrients (Ah-Fong et al., 2019), such as carbohydrates and lipids. For metabolites for which changes in abundance are measured, uptake and demand (transport) reactions should be added to the model.

Fourth, the relative pathogen biomass composition and growth rates should be measured. A promising approach for this is Fourier-transform infrared spectroscopy (FTIR; Mayers et al., 2013). This method was optimized for analyses of brown algae and was successfully used in the reconstruction of a GEM for the diatom *Phaeodactylum tricornutum* to quantify the percentages of carbohydrate, protein, DNA/RNA, and fatty acids per gram of cellular dry weight (Levering et al., 2016). Once the main classes of biomass components are quantified and are related to growth rate, more specific metabolites can be assigned based on traditional metabolomics methods (e.g., chromatography and mass spectrometry) and incorporated into the GEM.

Fifth, predicted phenotypes, e.g., induced by specific nutrient starvation, should be validated by growth experiments and gene/reaction essentiality should be validated in knockout or knockdown mutants, for instance, as was done for the nitrate assimilation cluster (Abrahamian et al., 2016). As discussed, we anticipate that CRISPR-Cas gene editing will be successfully employed in many pathogens in the coming years, but as an alternative, gene silencing mutants can be generated to investigate metabolic perturbations. The model should be updated with novel findings, and discrepancies should lead to corrections of the model. For instance, when the knockout of a predicted essential gene is not lethal *in vitro*, there are likely alternative enzymes or metabolic routes that compensate for this mutation. The model should be inspected on incorrect annotations or missing reactions accordingly.

Sixth, the refined pathogen GEM should be integrated with a similarly refined GEM for its host. This necessitates deploying more sophisticated constraints and objective functions to simulate a more realistic symbiosis for this pathosystem, such as multi-objective simulations to address the competition for nutrients (Jamshidi and Raghunathan, 2015).

In addition to the here proposed steps, there is a large and rapidly increasing number of methods and algorithms that can be applied to GEMs to gain further insights into the complex system of pathogen–host interactions (Lewis et al., 2012). For instance, regulatory networks could be inferred from (anti-)correlated expression patterns in dual RNA-Seq data and other experimental data and integrated into GEMs to further constrain the fluxes, in order to learn how metabolism is regulated during infection (Peyraud et al., 2018). Collectively, we anticipate that these steps will lead to high-quality pathogen and pathogen–host GEMs that can be used to identify novel targets for disease control and further help to understand how pathogens interact with their hosts.

CONCLUSION

Systems biology, in particular GEMs, offers a unique approach to study oomycetes and their intricate interactions with their hosts. GEMs not only offer a holistic overview of metabolism, but also constitute a foundation on which to incorporate omics measurements at various levels, allowing integrated analyses of key processes in pathogenesis and pathogen–host interaction. Although there are still a number of remaining technical and methodological challenges, GEMs hold great promise for providing mechanistic insight into strategies exploited by oomycetes to proliferate and infect their hosts, ultimately allowing us to develop new means of controlling these highly relevant pathogens.

AUTHOR CONTRIBUTIONS

SR wrote the first draft of the manuscript. FG edited sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was funded by the Food-for-Thought campaign from Wageningen University Fund.

REFERENCES

Abdel-Haleem, A. M., Hefzi, H., Mineta, K., Gao, X., Gojobori, T., Palsson, B. O., et al. (2018). Functional interrogation of *plasmodium* genus metabolism identifies species-and stage-specific differences in nutrient essentiality and drug targeting. *PLoS Comput. Biol.* 14:e1005895. doi: 10.1371/journal.pcbi.1005895

Abrahamian, M., Ah-Fong, A. M. V., Davis, C., Andreeva, K., and Judelson, H. S. (2016). Gene expression and silencing studies in *Phytophthora infestans* reveal infection-specific nutrient transporters and a role for the nitrate reductase pathway in plant pathogenesis. *PLoS Pathog.* 12:e1006097. doi: 10.1371/journal.ppat.1006097

Abrahamian, M., Kagda, M., Ah-Fong, A. M. V., and Judelson, H. S. (2017). Rethinking the evolution of eukaryotic metabolism: novel cellular partitioning of enzymes in stramenopiles links serine biosynthesis to glycolysis in mitochondria. *BMC Evol. Biol.* 17, 1–16. doi: 10.1186/s12862-017-1087-8

Aderem, A. (2005). Systems biology: its practice and challenges. *Cell* 121, 511–513. doi: 10.1016/j.cell.2005.04.020

Agren, R., Bordel, S., Mardinoglu, A., Pornputtapong, N., Nookaew, I., and Nielsen, J. (2012). Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using init. *PLoS Comput. Biol.* 8:e1002518. doi: 10.1371/journal.pcbi.1002518

Ah-Fong, A. M. V., Boyd, A. M., Matson, M. E. H., and Judelson, H. S. (2021). A Cas12a-based gene editing system for *Phytophthora infestans* reveals monoallelic expression of an elicitor. *Mol. Plant Pathol.* 22, 737–752. doi: 10.1111/mpp.13051

Ah-Fong, A. M. V., Kagda, M. S., Abrahamian, M., and Judelson, H. S. (2019). Niche-specific metabolic adaptation in biotrophic and necrotrophic oomycetes is manifested in differential use of nutrients, variation in gene content, and enzyme evolution. *PLoS Pathog.* 15:e1007729. doi: 10.1371/journal.ppat.1007729

Ah-Fong, A. M. V., Kim, K. S., and Judelson, H. S. (2017a). RNA-seq of life stages of the oomycete *Phytophthora infestans* reveals dynamic changes in metabolic, signal transduction, and pathogenesis genes and a major role for calcium signaling in development. *BMC Genomics* 18, 1–21. doi: 10.1186/s12864-017-3585-x

Ah-Fong, A. M. V., Shrivastava, J., and Judelson, H. S. (2017b). Lifestyle, gene gain and loss, and transcriptional remodeling cause divergence in the transcriptomes of *Phytophthora infestans* and *Pythium ultimum* during potato tuber colonization. *BMC Genomics* 18, 1–28. doi: 10.1186/s12864-017-4151-2

Albert, R. (2007). Network inference, analysis, and modeling in systems biology. *Plant Cell* 19, 3327–3338. doi: 10.1105/tpc.107.054700

Andronis, C. E., Hane, J. K., Bringans, S., Hardy, G. E. S., Jacques, S., Lipscombe, R., et al. (2020). Gene validation and remodelling using proteogenomics of *Phytophthora cinnamomi*, the causal agent of dieback. *Front. Microbiol.* 12:665396. doi: 10.3389/fmicb.2021.665396

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., et al. (2019). KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 36, 2251–2252. doi: 10.1101/602110

Bartocci, E., and Lió, P. (2016). Computational modeling, formal analysis, and tools for systems biology. *PLoS Comput. Biol.* 12:e1004591. doi: 10.1371/journal.pcbi.1004591

Baxter, L., Tripathy, S., Ishaque, N., Boot, N., Cabral, A., Kemen, E., et al. (2010). Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* 330, 1549–1551. doi: 10.1126/science.1195203

Bazzani, S., Hoppe, A., and Holzhütter, H.-G. (2012). Network-based assessment of the selectivity of metabolic drug targets in *plasmodium falciparum* with respect to human liver metabolism. *BMC Syst. Biol.* 6:118. doi: 10.1186/1752-0509-6-118

Beakes, G. W., Glockling, S. L., and Sekimoto, S. (2011). The evolutionary phylogeny of the oomycete fungi. *Protoplasma* 249, 3–19. doi: 10.1007/s00709-011-0269-2

Boevink, P. C., Birch, P. R. J., Turnbull, D., and Whisson, S. C. (2020). Devastating intimacy: the cell biology of plant–*Phytophthora* interactions. *New Phytol.* 228, 445–458. doi: 10.1111/nph.16650

Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B. Ø., and Jamshidi, N. (2010). Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Mol. Syst. Biol.* 6:422. doi: 10.1038/msb.2010.68

Bordbar, A., Monk, J. M., King, Z. A., and Palsson, B. O. (2014). Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* 15, 107–120. doi: 10.1038/nrg3643

Botero, K., Restrepo, S., and Pinzón, A. (2018b). A genome-scale metabolic model of potato late blight suggests a photosynthesis suppression mechanism. *BMC Genomics* 19, 31–44. doi: 10.1186/s12864-018-5192-x

Botero, D., Valdés, I., Rodríguez, M.-J., Henao, D., Danies, G., González, A. F., et al. (2018a). A genome-scale metabolic reconstruction of Phytophthora infestans with the integration of transcriptional data reveals the key metabolic patterns involved in the interaction of its host. *Front. Genet.* 9:244. doi: 10.3389/fgene.2018.00244

Breitling, R. (2010). What is systems biology? *Front. Physiol.* 1, 1662–1664. doi: 10.3389/fphys.2010.00009

Bronkhorst, J., Kasteel, M., van Veen, S., Clough, J. M., Kots, K., Buijs, J., et al. (2021). A slicing mechanism facilitates host entry by plant-pathogenic *Phytophthora*. *Nat. Microbiol.* 6, 1000–1006. doi: 10.1038/s41564-021-00919-7

Carey, M. A., Papin, J. A., and Guler, J. L. (2017). Novel *plasmodium falciparum* metabolic network reconstruction identifies shifts associated with clinical antimalarial resistance. *BMC Genomics* 18, 1–19. doi: 10.1101/119941

Castro, J. C., Valdés, I., Gonzalez-García, L. N., Danies, G., Cañas, S., Winck, F. V., et al. (2019). Gene regulatory networks on transfer entropy (GRNTE): a novel approach to reconstruct gene regulatory interactions applied to a case study for the plant pathogen *Phytophthora infestans*. *Theor. Biol. Med. Model.* 16, 1–15. doi: 10.1186/s12976-019-0103-7

Cesur, M. F., Abdik, E., Güven-Gülhan, Ü., Durmuş, S., and Çakır, T. (2018). "Computational systems biology of metabolism in infection," in *Metabolic Interaction in Infection. Experientia Supplementum*. eds. R. Silvestre and E. Torrado (Springer, Cham) 109, 235–283.

Chauhan, N., and Singh, S. (2019). Integrative computational framework for understanding metabolic modulation in *Leishmania*. *Front. Bioeng. Biotechnol.* 7:336. doi: 10.3389/fbioe.2019.00336

Chavali, A. K., D'Auria, K. M., Hewlett, E. L., Pearson, R. D., and Papin, J. A. (2012). A metabolic network approach for the identification and prioritization of antimicrobial drug targets. *Trends Microbiol.* 20, 113–123. doi: 10.1016/j.tim.2011.12.004

Chen, L. (2013). SWEET sugar transporters for phloem transport and pathogen nutrition. *New Phytol.* 201, 1150–1155. doi: 10.1111/nph.12445

Cook, D. E., Mesarich, C. H., and Thomma, B. P. H. J. (2015). Understanding plant immunity as a surveillance system to detect invasion. *Annu. Rev. Phytopathol.* 53, 541–563. doi: 10.1146/annurev-phyto-080614-120114

De Bary, A. (1876). Researches into the nature of the potato fungus, *Phytophthora infestans*. *J. R. Agric. Soc. Engl.* 12, 239–269.

DeBerardinis, R. J., and Thompson, C. B. (2012). Cellular metabolism and disease: what do metabolic outliers teach us? *Cell* 148, 1132–1144. doi: 10.1016/j.cell.2012.02.032

Denton, J. F., Lugo-Martinez, J., Tucker, A. E., Schrider, D. R., Warren, W. C., and Hahn, M. W. (2014). Extensive error in the number of genes inferred from draft genome assemblies. *PLoS Comput. Biol.* 10:e1003998. doi: 10.1371/journal.pcbi.1003998

Derevnina, L., Dagdas, Y. F., De la Concepcion, J. C., Bialas, A., Kellner, R., Petre, B., et al. (2016a). Nine things to know about elicitins. *New Phytol.* 212, 888–895. doi: 10.1111/nph.14137

Derevnina, L., Petre, B., Kellner, R., Dagdas, Y. F., Sarowar, M. N., Giannakopoulou, A., et al. (2016b). Emerging oomycete threats to plants and animals. *Philos. Trans. Royal. Soc. B: Biol. Sci.* 371:20150459. doi: 10.1098/rstb.2015.0459

Dix, A., Vlaic, S., Guthke, R., and Linde, J. (2016). Use of systems biology to decipher host–pathogen interaction networks and predict biomarkers. *Clin. Microbiol. Infect.* 22, 600–606. doi: 10.1016/j.cmi.2016.04.014

Dunphy, L. J., and Papin, J. A. (2018). Biomedical applications of genome-scale metabolic network reconstructions of human pathogens. *Curr. Opin. Biotechnol.* 51, 70–79. doi: 10.1016/j.copbio.2017.11.014

Durmuş, S., Çakir, T., Özgür, A., and Guthke, R. (2015). A review on computational systems biology of pathogen-host interactions. *Front. Microbiol.* 6:235. doi: 10.3389/fmicb.2015.00235

Durmuş Tekir, S. D., and Ülgen, K. Ö. (2012). Systems biology of pathogen-host interaction: networks of protein-protein interaction within pathogens and pathogen-human interactions in the post-genomic era. *Biotechnol. J.* 8, 85–96. doi: 10.1002/biot.201200110

Eberl, G. (2018). Robustness in living organisms is homeostasis. *Semin. Immunol.* 36, 56–57. doi: 10.1016/j.smim.2017.12.007

Edwards, J. S., and Palsson, B. O. (1999). Systems properties of the *haemophilus influenzae* Rd metabolic genotype. *J. Biol. Chem.* 274, 17410–17416. doi: 10.1074/jbc.274.25.17410

Faino, L., Seidl, M. F., Datema, E., van den Berg, G. C. M., Janssen, A., Wittenberg, A. H. J., et al. (2015). Single-molecule real-time sequencing combined with optical mapping yields completely finished fungal genome. *MBio* 6, e00936–e00915. doi: 10.1128/mbio.00936-15

Fang, Y., and Tyler, B. M. (2015). Efficient disruption and replacement of an effector gene in the oomycete *Phytophthora sojae* using CRISPR/Cas9. *Mol. Plant Pathol.* 17, 127–139. doi: 10.1101/025023

Fawke, S., Doumane, M., and Schornack, S. (2015). Oomycete interactions with plants: infection strategies and resistance principles. *Microbiol. Mol. Biol. Rev.* 79, 263–280. doi: 10.1128/MMBR.00010-15

Feist, A. M., and Palsson, B. O. (2010). The biomass objective function. *Curr. Opin. Microbiol.* 13, 344–349. doi: 10.1016/j.mib.2010.03.003

Fernandes, B. S., Dias, O., Costa, G., Kaupert Neto, A. A., Resende, T. F. C., Oliveira, J. V. C., et al. (2019). Genome-wide sequencing and metabolic annotation of *Pythium irregulare* CBS 494.86: understanding eicosapentaenoic acid production. *BMC Biotechnol.* 19, 1–17. doi: 10.1186/s12896-019-0529-3

Fletcher, K., Gil, J., Bertier, L. D., Kenefick, A., Wood, K. J., Zhang, L., et al. (2019). Genomic signatures of heterokaryosis in the oomycete pathogen *Bremia lactucae*. *Nat. Commun.* 10, 1–13. doi: 10.1038/s41467-019-10550-0

Fletcher, K., Klosterman, S. J., Derevnina, L., Martin, F., Bertier, L. D., Koike, S., et al. (2018). Comparative genomics of downy mildews reveals potential adaptations to biotrophy. *BMC Genomics* 19, 1–23. doi: 10.1186/s12864-018-5214-8

Frantzeskakis, L., Di Pietro, A., Rep, M., Schirawski, J., Wu, C., and Panstruga, R. (2019). Rapid evolution in plant–microbe interactions – a molecular genomics perspective. *New Phytol.* 225, 1134–1142. doi: 10.1111/nph.15966

García Sánchez, C. E., and Torres Sáez, R. G. (2014). Comparison and analysis of objective functions in flux balance analysis. *Biotechnol. Prog.* 30, 985–991. doi: 10.1002/btpr.1949

Grenville-Briggs, L. J., Anderson, V. L., Fugelstad, J., Avrova, A. O., Bouzenzana, J., Williams, A., et al. (2008). Cellulose synthesis in *Phytophthora infestans* is required for normal appressorium formation and successful infection of potato. *Plant Cell* 20, 720–738. doi: 10.1105/tpc.107.052043

Griffiths, R. G., Dancer, J., O'Neill, E., and Harwood, J. L. (2003). Effect of culture conditions on the lipid composition of *Phytophthora infestans*. *New Phytol.* 158, 337–344. doi: 10.1046/j.1469-8137.2003.00738.x

Gu, C., Kim, G. B., Kim, W. J., Kim, H. U., and Lee, S. Y. (2019). Current status and applications of genome-scale metabolic models. *Genome Biol.* 20:121. doi: 10.1186/s13059-019-1730-3

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., et al. (2009). Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 461, 393–398. doi: 10.1038/nature08358

Hodgson, W. A. (1958). Growth of four races of *Phytophthora infestans* (Mont.) De Bary in synthetic media. *Can. J. Plant Sci.* 38, 145–154. doi: 10.4141/cjps58-026

Hohl, H. R. (1991). "Nutrition," in *Phytophthora Infestans, the Cause of Late Blight of Potato*. eds. D. S. Ingram and P. H. Williams (Academic Press, London: Advances in Plant Pathology), 53–83.

Horn, F., Heinekamp, T., Kniemeyer, O., Pollmächer, J., Valiante, V., and Brakhage, A. A. (2012). Systems biology of fungal infection. *Front. Microbiol.* 3:108. doi: 10.3389/fmicb.2012.00108

Huthmacher, C., Hoppe, A., Bulik, S., and Holzhütter, H.-G. (2010). Antimalarial drug targets in *plasmodium falciparum* predicted by stage-specific metabolic network analysis. *BMC Syst. Biol.* 4, 1–27. doi: 10.1186/1752-0509-4-120

Hyduke, D. R., Lewis, N. E., and Palsson, B. Ø. (2013). Analysis of omics data with genome-scale models of metabolism. *Mol. BioSyst.* 9, 167–174. doi: 10.1039/C2MB25453K

Ideker, T., Galitski, T., and Hood, L. (2001). A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* 2, 343–372. doi: 10.1146/annurev.genom.2.1.343

Jamshidi, N., and Raghunathan, A. (2015). Cell scale host-pathogen modeling: another branch in the evolution of constraint-based methods. *Front. Microbiol.* 6:1032. doi: 10.3389/fmicb.2015.01032

Judelson, H. S. (2017). Metabolic diversity and novelties in the oomycetes. *Annu. Rev. Microbiol.* 71, 21–39. doi: 10.1146/annurev-micro-090816-093609

Judelson, H. S., and Ah-Fong, A. M. V. (2018). Exchanges at the plant-oomycete interface that influence disease. *Plant Physiol.* 179, 1198–1211. doi: 10.1104/pp.18.00979

Kamoun, S., Furzer, O., Jones, J. D. G., Judelson, H. S., Ali, G. S., Dalio, R. J. D., et al. (2014). The top 10 oomycete pathogens in molecular plant pathology. *Mol. Plant Pathol.* 16, 413–434. doi: 10.1111/mpp.12190

Kavvas, E. S., Seif, Y., Yurkovich, J. T., Norsigian, C., Poudel, S., Greenwald, W. W., et al. (2018). Updated and standardized genome-scale reconstruction of *mycobacterium tuberculosis* H37Rv, iEK1011, simulates flux states indicative of physiological conditions. *BMC Syst. Biol.* 12, 1–15. doi: 10.1186/s12918-018-0557-y

Keeling, P. J., and Burki, F. (2019). Progress towards the tree of eukaryotes. *Curr. Biol.* 29, R808–R817. doi: 10.1016/j.cub.2019.07.031

Kemen, E., and Jones, J. D. G. (2012). Obligate biotroph parasitism: can we link genomes to lifestyles? *Trends Plant Sci.* 17, 448–457. doi: 10.1016/j.tplants.2012.04.005

Kitano, H. (2002). Systems biology: a brief overview. *Science* 295, 1662–1664. doi: 10.1126/science.1069492

Kloehn, J., Blume, M., Cobbold, S., Saunders, E., Dagley, M., and McConville, M. (2016). Using metabolomics to dissect host–parasite interactions. *Curr. Opin. Microbiol.* 32, 59–65. doi: 10.1016/j.mib.2016.04.019

Koduru, L., Kim, H. Y., Lakshmanan, M., Mohanty, B., Lee, Y. Q., Lee, C. H., et al. (2020). Genome-scale metabolic reconstruction and in silico analysis of the rice leaf blight pathogen, *Xanthomonas oryzae*. *Mol. Plant Pathol.* 21, 527–540. doi: 10.1111/mpp.12914

Lazar, M. A., and Birnbaum, M. J. (2012). De-meaning of metabolism. *Science* 336, 1651–1652. doi: 10.1126/science.1221834

Leesutthiphonchai, W., Vu, A. L., Ah-Fong, A. M. V., and Judelson, H. S. (2018). How does *Phytophthora infestans* evade control efforts? Modern insight into the late blight disease. *Phytopathology* 108, 916–924. doi: 10.1094/PHYTO-04-18-0130-IA

Levering, J., Broddrick, J., Dupont, C. L., Peers, G., Beeri, K., Mayers, J., et al. (2016). Genome-scale model reveals metabolic basis of biomass partitioning in a model diatom. *PLoS One* 11:e0155038. doi: 10.1371/journal.pone.0155038

Lewis, N. E., Nagarajan, H., and Palsson, B. O. (2012). Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods. *Nat. Rev. Microbiol.* 10, 291–305. doi: 10.1038/nrmicro2737

López-Calcagno, P. E., Moreno, J., Cedeño, L., Labrador, L., Concepción, J. L., and Avilán, L. (2009). Cloning, expression and biochemical characterization of mitochondrial and cytosolic malate dehydrogenase from *Phytophthora infestans*. *Mycol. Res.* 113, 771–781. doi: 10.1016/j.mycres.2009.02.012

Malar, C. M., Yuzon, J. D., Das, S., Das, A., Panda, A., Ghosh, S., et al. (2019). Haplotype-phased genome assembly of virulent *Phytophthora ramorum* isolate ND886 facilitated by long-read sequencing reveals effector polymorphisms and copy number variation. *Mol. Plant-Microbe Interact.* 32, 1047–1060. doi: 10.1094/MPMI-08-18-0222-R

Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787–3793. doi: 10.1093/bioinformatics/bti430

Mayers, J. J., Flynn, K. J., and Shields, R. J. (2013). Rapid determination of bulk microalgal biochemical composition by Fourier-transform infrared spectroscopy. *Bioresour. Technol.* 148, 215–220. doi: 10.1016/j.biortech.2013.08.133

McDowell, J. M. (2011). Genomes of obligate plant pathogens reveal adaptations for obligate parasitism. *Proc. Natl. Acad. Sci.* 108, 8921–8922. doi: 10.1073/pnas.1105802108

McGowan, J., Byrne, K. P., and Fitzpatrick, D. A. (2018). Comparative analysis of oomycete genome evolution using the oomycete gene order browser (OGOB). *Genome Biol. Evol.* 11, 189–206. doi: 10.1093/gbe/evy267

McGowan, J., and Fitzpatrick, D. A. (2017). Genomic, network, and phylogenetic analysis of the oomycete effector arsenal. *mSphere* 2, e00408–e00417. doi: 10.1128/msphere.00408-17

McGowan, J., and Fitzpatrick, D. A. (2020). Recent advances in oomycete genomics. *Adv. Genet.* 105, 175–228. doi: 10.1016/bs.adgen.2020.03.001

McKnight, S. L. (2010). On getting there from here. *Science* 330, 1338–1339. doi: 10.1126/science.1199908

Monk, J., Nogales, J., and Palsson, B. O. (2014). Optimizing genome-scale network reconstructions. *Nat. Biotechnol.* 32, 447–452. doi: 10.1038/nbt.2870

Mukhtar, M. S., Carvunis, A.-R., Dreze, M., Epple, P., Steinbrenner, J., Moore, J., et al. (2011). Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science* 333, 596–601. doi: 10.1126/science.1203659

Nagarajan, N., and Pop, M. (2013). Sequence assembly demystified. *Nat. Rev. Genet.* 14, 157–167. doi: 10.1038/nrg3367

Nakahigashi, K., Toya, Y., Ishii, N., Soga, T., Hasegawa, M., Watanabe, H., et al. (2009). Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Mol. Syst. Biol.* 5:306. doi: 10.1038/msb.2009.65

Nijhout, H. F., Best, J. A., and Reed, M. C. (2018). Systems biology of robustness and homeostatic mechanisms. *WIREs Syst. Biol. Med.* 11:e1440. doi: 10.1002/wsbm.1440

O'Brien, E. J., Monk, J. M., and Palsson, B. O. (2015). Using genome-scale models to predict biological capabilities. *Cell* 161, 971–987. doi: 10.1016/j.cell.2015.05.019

O'Hara, J. K., Kerwin, L. J., Cobbold, S. A., Tai, J., Bedell, T. A., Reider, P. J., et al. (2014). Targeting NAD+ metabolism in the human malaria parasite *plasmodium falciparum*. *PLoS One* 9:e94061. doi: 10.1371/journal.pone.0094061

Olive, A. J., and Sassetti, C. M. (2016). Metabolic crosstalk between host and pathogen: sensing, adapting and competing. *Nat. Rev. Microbiol.* 14, 221–234. doi: 10.1038/nrmicro.2016.12

Olszewski, K. L., Morrisey, J. M., Wilinski, D., Burns, J. M., Vaidya, A. B., Rabinowitz, J. D., et al. (2009). Host-parasite interactions revealed by *plasmodium falciparum* metabolomics. *Cell Host Microbe* 5, 191–199. doi: 10.1016/j.chom.2009.01.004

Orth, J. D., Thiele, I., and Palsson, B. Ø. (2010). What is flux balance analysis? *Nat. Biotechnol.* 28, 245–248. doi: 10.1038/nbt.1614

Pearson, W. R. (2013). An introduction to sequence similarity ("homology") searching. *Curr. Protoc. Bioinform.* 42, 3.1.1–3.1.8. doi: 10.1002/0471250953.bi0301s42

Perrine-Walker, F. (2020). *Phytophthora palmivora*–cocoa interaction. *J. Fungi.* 6:167. doi: 10.3390/jof6030167

Petre, B., Contreras, M. P., Bozkurt, T. O., Schattat, M. H., Sklenar, J., Schornack, S., et al. (2021). Host-interactor screens of *Phytophthora infestans* RXLR proteins reveal vesicle trafficking as a major effector-targeted process. *Plant Cell* 33, 1447–1471. doi: 10.1093/plcell/koab069

Pettongkhao, S., Navet, N., Schornack, S., Tian, M., and Churngchow, N. (2020). A secreted protein of 15 kDa plays an important role in *Phytophthora palmivora* development and pathogenicity. *Sci. Rep.* 10, 1–15. doi: 10.1038/s41598-020-59007-1

Peyraud, R., Cottret, L., Marmiesse, L., and Genin, S. (2018). Control of primary metabolism by a virulence regulatory network promotes robustness in a plant pathogen. *Nat. Commun.* 9, 1–14. doi: 10.1038/s41467-017-02660-4

Peyraud, R., Cottret, L., Marmiesse, L., Gouzy, J., and Genin, S. (2016). A resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen *Ralstonia solanacearum*. *PLoS Pathog.* 12:e1005939. doi: 10.1371/journal.ppat.1005939

Peyraud, R., Dubiella, U., Barbacci, A., Genin, S., Raffaele, S., and Roby, D. (2017). Advances on plant-pathogen interactions from molecular toward systems biology perspectives. *Plant J.* 90, 720–737. doi: 10.1111/tpj.13429

Peyraud, R., Mbengue, M., Barbacci, A., and Raffaele, S. (2019). Intercellular cooperation in a fungal plant pathogen facilitates host colonization. *Proc. Natl. Acad. Sci.* 116, 3193–3201. doi: 10.1073/pnas.1811267116

Pinzón, A., Barreto, E., Bernal, A., Achenie, L., González Barrios, A. F., Isea, R., et al. (2009). Computational models in plant-pathogen interactions: the case of *Phytophthora infestans*. *Theor. Biol. Med. Model.* 6, 1–11. doi: 10.1186/1742-4682-6-24

Pinzón, A., Rodriguez-R, L. M., Gonzalez, A., Bernal, A., and Restrepo, S. (2010). Targeted metabolic reconstruction: a novel approach for the characterization of plant-pathogen interactions. *Brief. Bioinform.* 12, 151–162. doi: 10.1093/bib/bbq009

Plata, G., Hsiao, T., Olszewski, K. L., Llinás, M., and Vitkup, D. (2010). Reconstruction and flux-balance analysis of the *plasmodium falciparum* metabolic network. *Mol. Syst. Biol.* 6:408. doi: 10.1038/msb.2010.60

Pratapa, A., Balachandran, S., and Raman, K. (2015). Fast-SL: an efficient algorithm to identify synthetic lethal sets in metabolic networks. *Bioinformatics* 31, 3299–3305. doi: 10.1093/bioinformatics/btv352

Prigent, S., Collet, G., Dittami, S. M., Delage, L., Ethis de Corny, F., Dameron, O., et al. (2014). The genome-scale metabolic network of *Ectocarpus siliculosus* (EctoGEM): a resource to study brown algal physiology and beyond. *Plant J.* 80, 367–381. doi: 10.1111/tpj.12627

Pritchard, L., and Birch, P. (2011). A systems biology perspective on plant–microbe interactions: biochemical and structural targets of pathogen effectors. *Plant Sci.* 180, 584–603. doi: 10.1016/j.plantsci.2010.12.008

Reed, J. L., Famili, I., Thiele, I., and Palsson, B. O. (2006). Towards multidimensional genome annotation. *Nat. Rev. Genet.* 7, 130–141. doi: 10.1038/nrg1769

Richards, T. A., Soanes, D. M., Jones, M. D. M., Vasieva, O., Leonard, G., Paszkiewicz, K., et al. (2011). Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. *Proc. Natl. Acad. Sci.* 108, 15258–15263. doi: 10.1073/pnas.1105100108

Richards, T. A., and Talbot, N. J. (2013). Horizontal gene transfer in osmotrophs: playing with public goods. *Nat. Rev. Microbiol.* 11, 720–727. doi: 10.1038/nrmicro3108

Rienksma, R. A., Schaap, P. J., Martins dos Santos, V. A. P., and Suarez-Diez, M. (2018). Modeling the metabolic state of *mycobacterium tuberculosis* upon infection. *Front. Cell. Infect. Microbiol.* 8:264. doi: 10.3389/fcimb.2018.00264

Rödelsperger, C., Athanasouli, M., Lenuzzi, M., Theska, T., Sun, S., Dardiry, M., et al. (2019). Crowdsourcing and the feasibility of manual gene annotation: a pilot study in the nematode *Pristionchus pacificus*. *Sci. Rep.* 9, 1–9. doi: 10.1038/s41598-019-55359-5

Rodenburg, S. Y. A., de Ridder, D., Govers, F., and Seidl, M. F. (2020). Oomycete metabolism is highly dynamic and reflects lifestyle adaptations. bioRxiv. [Preprint] doi: 10.1101/2020.02.12.941195

Rodenburg, S. Y. A., Seidl, M. F., de Ridder, D., and Govers, F. (2017). Genome-wide characterization of Phytophthora infestans metabolism: a systems biology approach. *Mol. Plant Pathol.* 19, 1403–1413. doi: 10.1101/171082

Rodenburg, S. Y. A., Seidl, M. F., Judelson, H. S., Vu, A. L., Govers, F., and de Ridder, D. (2019). Metabolic model of the Phytophthora infestans-tomato interaction reveals metabolic switches during host colonization. *MBio* 10, e00419–e00454. doi: 10.1128/mbio.00454-19

Sahoo, S., Aurich, M. K., Jonsson, J. J., and Thiele, I. (2014). Membrane transporters in a human genome-scale metabolic knowledgebase and their implications for disease. *Front. Physiol.* 5:91. doi: 10.3389/fphys.2014.00091

Salzberg, S. L. (2019). Next-generation genome annotation: we still struggle to get it right. *Genome Biol.* 20:92. doi: 10.1186/s13059-019-1715-2

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nat. Ecol. Evol.* 3, 430–439. doi: 10.1038/s41559-018-0793-y

Savory, F. R., Milner, D. S., Miles, D. C., and Richards, T. A. (2018). Ancestral function and diversification of a horizontally acquired oomycete carboxylic acid transporter. *Mol. Biol. Evol.* 35, 1887–1900. doi: 10.1093/molbev/msy082

Seidl, M. F., Schneider, A., Govers, F., and Snel, B. (2013). A predicted functional gene network for the plant pathogen *Phytophthora infestans* as a framework for genomic biology. *BMC Genomics* 14:483. doi: 10.1186/1471-2164-14-483

Seidl, M. F., Van den Ackerveken, G., Govers, F., and Snel, B. (2010). A domain-centric analysis of oomycete plant pathogen genomes reveals unique protein organization. *Plant Physiol.* 155, 628–644. doi: 10.1104/pp.110.167841

Seppey, M., Manni, M., and Zdobnov, E. M. (2019). "BUSCO: Assessing genome assembly and annotation completeness," in *Gene Prediction. Methods in Molecular Biology*. ed. M. Kollmar (Humana, New York, NY), 1962, 227–245.

Sharma, M., Shaikh, N., Yadav, S., Singh, S., and Garg, P. (2017). A systematic reconstruction and constraint-based analysis of *Leishmania donovani* metabolic network: identification of potential antileishmanial drug targets. *Mol. BioSyst.* 13, 955–969. doi: 10.1039/C6MB00823B

Stajich, J. E., Vu, A. L., Judelson, H. S., Vogel, G. M., Gore, M. A., Carlson, M. O., et al. (2021). High-quality reference genome sequence for the oomycete vegetable pathogen *Phytophthora capsici* strain LT1534. *Microbiol. Res. Announce.* 10, e00295–e00221. doi: 10.1128/mra.00295-21

Stanway, R. R., Bushell, E., Chiappino-Pepe, A., Roques, M., Sanderson, T., Franke-Fayard, B., et al. (2019). Genome-scale identification of essential metabolic processes for targeting the *plasmodium* liver stage. *Cell* 179, 1112–1128. doi: 10.1016/j.cell.2019.10.030

Subramanian, A., Jhawar, J., and Sarkar, R. R. (2015). Dissecting *Leishmania infantum* energy metabolism - a systems perspective. *PLoS One* 10:e0137976. doi: 10.1371/journal.pone.0137976

Sun, Q., Liu, J., Zhang, Q., Qing, X., Dobson, G., Li, X., et al. (2012). Characterization of three novel desaturases involved in the delta-6 desaturation pathways for polyunsaturated fatty acid biosynthesis from *Phytophthora infestans*. *Appl. Microbiol. Biotechnol.* 97, 7689–7697. doi: 10.1007/s00253-012-4613-z

Thiele, I., and Palsson, B. Ø. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* 5, 93–121. doi: 10.1038/nprot.2009.203

Thines, M., Sharma, R., Rodenburg, S. Y. A., Gogleva, A., Judelson, H. S., Xia, X., et al. (2020). The genome of *Peronospora belbahrii* reveals high heterozygosity, a low number of canonical effectors, and TC-rich promoters. *Mol. Plant-Microbe Interact.* 33, 742–753. doi: 10.1094/MPMI-07-19-0211-R

Thomma, B. P. H. J., Seidl, M. F., Shi-Kunne, X., Cook, D. E., Bolton, M. D., van Kan, J. A. L., et al. (2016). Mind the gap; seven reasons to close fragmented genome assemblies. *Fungal Genet. Biol.* 90, 24–30. doi: 10.1016/j.fgb.2015.08.010

Turner, R. S. (2005). After the famine: plant pathology, *Phytophthora infestans*, and the late blight of potatoes, 1845–1960. *Hist. Stud. Phys. Biol. Sci.* 35, 341–370. doi: 10.1525/hsps.2005.35.2.341

Turner, R. S. (2008). Potato agriculture, late blight science, and the molecularization of plant pathology. *Hist. Stud. Nat. Sci.* 38, 223–257. doi: 10.1525/hsns.2008.38.2.223

Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R. H. Y., Aerts, A., et al. (2006). *Phytophthora genome* sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313, 1261–1266. doi: 10.1126/science.1128796

Tymoshenko, S., Oppenheim, R. D., Agren, R., Nielsen, J., Soldati-Favre, D., and Hatzimanikatis, V. (2015). Metabolic needs and capabilities of *toxoplasma gondii* through combined computational and experimental analysis. *PLoS Comput. Biol.* 11:e1004261. doi: 10.1371/journal.pcbi.1004261

van den Hoogen, J., and Govers, F. (2018a). Attempts to implement CRISPR/Cas9 for genome editing in the oomycete *Phytophthora infestans*. bioRxiv:274829. [Preprint] doi: 10.1101/274829

van den Hoogen, J., and Govers, F. (2018b). GPCR-bigrams: enigmatic signaling components in oomycetes. *PLoS Pathog.* 14:e1007064. doi: 10.1371/journal.ppat.1007064

Wang, S., Boevink, P. C., Welsh, L., Zhang, R., Whisson, S. C., and Birch, P. R. J. (2017). Delivery of cytoplasmic and apoplastic effectors from *Phytophthora infestans* haustoria by distinct secretion pathways. *New Phytol.* 216, 205–215. doi: 10.1111/nph.14696

Wang, W., Liu, X., and Govers, F. (2021). The mysterious route of sterols in oomycetes. *PLoS Pathog.* 17:e1009591. doi: 10.1371/journal.ppat.1009591

Wang, Y., Tyler, B. M., and Wang, Y. (2019b). Defense and counterdefense during plant-pathogenic oomycete infection. *Annu. Rev. Microbiol.* 73, 667–696. doi: 10.1146/annurev-micro-020518-120022

Wang, S., Welsh, L., Thorpe, P., Whisson, S. C., Boevink, P. C., and Birch, P. R. J. (2018). The Phytophthora infestans haustorium is a site for secretion of diverse classes of infection-associated proteins. *MBio* 9, e01216–e01218. doi: 10.1128/mbio.01216-18

Wang, W., Xue, Z., Miao, J., Cai, M., Zhang, C., Li, T., et al. (2019a). PcMuORP1, an oxathiapiprolin-resistance gene, functions as a novel selection marker for *Phytophthora* transformation and CRISPR/Cas9 mediated genome editing. *Front. Microbiol.* 10:2402. doi: 10.3389/fmicb.2019.02402

Watson, E., Yilmaz, L. S., and Walhout, A. J. M. (2015). Understanding metabolic regulation at a systems level: metabolite sensing, mathematical predictions, and model organisms. *Annu. Rev. Genet.* 49, 553–575. doi: 10.1146/annurev-genet-112414-055257

Weßling, R., Epple, P., Altmann, S., He, Y., Yang, L., Henz, S. R., et al. (2014). Convergent targeting of a common host protein-network by pathogen effectors from three kingdoms of life. *Cell Host Microbe* 16, 364–375. doi: 10.1016/j.chom.2014.08.004

Whisson, S. C., Boevink, P. C., Wang, S., and Birch, P. R. (2016). The cell biology of late blight disease. *Curr. Opin. Microbiol.* 34, 127–135. doi: 10.1016/j.mib.2016.09.002

Whitaker, J. W., McConkey, G. A., and Westhead, D. R. (2009). The transferome of metabolic genes explored: analysis of the horizontal transfer of enzyme encoding genes in unicellular eukaryotes. *Genome Biol.* 10:R36. doi: 10.1186/gb-2009-10-4-r36

Wood, V., Lock, A., Harris, M. A., Rutherford, K., Bähler, J., and Oliver, S. G. (2019). Hidden in plain sight: what remains to be discovered in the eukaryotic proteome? *Open Biol.* 9:180241. doi: 10.1098/rsob.180241

Yandell, M., and Ence, D. (2012). A beginner's guide to eukaryotic genome annotation. *Nat. Rev. Genet.* 13, 329–342. doi: 10.1038/nrg3174

Yilmaz, L. S., and Walhout, A. J. (2017). Metabolic network modeling with model organisms. *Curr. Opin. Chem. Biol.* 36, 32–39. doi: 10.1016/j.cbpa.2016.12.025

Yuan, H., Cheung, C. Y. M., Poolman, M. G., Hilbers, P. A. J., and Riel, N. A. W. (2016). A genome-scale metabolic network reconstruction of tomato (*Solanum lycopersicum* L.) and its application to photorespiratory metabolism. *Plant J.* 85, 289–304. doi: 10.1111/tpj.13075

Yurkovich, J. T., and Palsson, B. O. (2016). Solving puzzles with missing pieces: the power of systems biology. *Proc. IEEE* 104, 2–7. doi: 10.1109/JPROC.2015.2505338

Zhang, C., and Hua, Q. (2016). Applications of genome-scale metabolic models in biotechnology and systems medicine. *Front. Physiol.* 6:413. doi: 10.3389/fphys.2015.00413

Zoledowska, S., Presta, L., Fondi, M., Decorosi, F., Giovannetti, L., Mengoni, A., et al. (2018). Metabolic modeling of Pectobacterium parmentieri SCC3193 provides insights into metabolic pathways of plant pathogenic bacteria. *Microorganisms 7*:101. doi: 10.1101/284968

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 16 November 2021
doi: 10.3389/fmicb.2021.765872

Dual-Transcriptomic, Microscopic, and Biocontrol Analyses of the Interaction Between the Bioeffector *Pythium oligandrum* and the Pythium Soft-Rot of Ginger Pathogen *Pythium myriotylum*

*Paul Daly[1†], Siqiao Chen[1,2†], Taiqiang Xue[3], Jingjing Li[3], Taha Majid Mahmood Sheikh[1], Qimeng Zhang[1], Xuehai Wang[4], Jinfeng Zhang[1], David A. Fitzpatrick[5], Jamie McGowan[5], Xiujuan Shi[4], Sheng Deng[1], Min Jiu[3], Dongmei Zhou[1], Irina S. Druzhinina[2] and Lihui Wei[1,6]**

[1] Key Lab of Food Quality and Safety of Jiangsu Province—State Key Laboratory Breeding Base, Institute of Plant Protection, Jiangsu Academy of Agricultural Sciences, Nanjing, China, [2] Fungal Genomics Laboratory (FungiG), Jiangsu Provincial Key Lab of Organic Solid Waste Utilization, Nanjing Agricultural University, Nanjing, China, [3] College of Food and Bioengineering, Henan University of Science and Technology, Luoyang, China, [4] Jinan Academy of Agricultural Sciences, Jinan, China, [5] Genome Evolution Laboratory, Maynooth University, Maynooth, Ireland, [6] School of Environment and Safety Engineering, Jiangsu University, Zhenjiang, China

OPEN ACCESS

Edited by:
Vaibhav Srivastava, Royal Institute of Technology, Sweden

Reviewed by:
Filippo De Curtis, Università degli Studi del Molise, Italy
Lauren Sara McKee, Royal Institute of Technology, Sweden

***Correspondence:**
Lihui Wei
weilihui@jaas.ac.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *27 August 2021*
Accepted: *20 October 2021*
Published: *16 November 2021*

Citation:
Daly P, Chen S, Xue T, Li J, Sheikh TMM, Zhang Q, Wang X, Zhang J, Fitzpatrick DA, McGowan J, Shi X, Deng S, Jiu M, Zhou D, Druzhinina IS and Wei L (2021) Dual-Transcriptomic, Microscopic, and Biocontrol Analyses of the Interaction Between the Bioeffector Pythium oligandrum and the Pythium Soft-Rot of Ginger Pathogen Pythium myriotylum. Front. Microbiol. 12:765872. doi: 10.3389/fmicb.2021.765872

Biological control is a promising approach to suppress diseases caused by *Pythium* spp. such as Pythium soft rot of ginger caused by *P. myriotylum*. Unusually for a single genus, it also includes species that can antagonize *Pythium* plant pathogens, such as *Pythium oligandrum*. We investigated if a new isolate of *P. oligandrum* could antagonize *P. myriotylum*, what changes occurred in gene expression when *P. oligandrum* (antagonist) and *P. myriotylum* (host) interacted, and whether *P. oligandrum* could control soft-rot of ginger caused by *P. myriotylum*. An isolate of *P. oligandrum*, GAQ1, recovered from soil could antagonize *P. myriotylum* in a plate-based confrontation assay whereby *P. myriotylum* became non-viable. The loss of viability of *P. myriotylum* coupled with how *P. oligandrum* hyphae could coil around and penetrate the hyphae of *P. myriotylum*, indicated a predatory interaction. We investigated the transcriptional responses of *P. myriotylum* and *P. oligandrum* using dual-RNAseq at a stage in the confrontation where similar levels of total transcripts were measured from each species. As part of the transcriptional response of *P. myriotylum* to the presence of *P. oligandrum*, genes including a subset of putative Kazal-type protease inhibitors were strongly upregulated along with cellulases, elicitin-like proteins and genes involved in the repair of DNA double-strand breaks. In *P. oligandrum*, proteases, cellulases, and peroxidases featured prominently in the upregulated genes. The upregulation along with constitutive expression of *P. oligandrum* proteases appeared to be responded to by the upregulation of putative protease inhibitors from *P. myriotylum*, suggesting a *P. myriotylum* defensive strategy. Notwithstanding this *P. myriotylum* defensive strategy, *P. oligandrum* had a

strong disease control effect on soft-rot of ginger caused by *P. myriotylum*. The newly isolated strain of *P. oligandrum* is a promising biocontrol agent for suppressing the soft-rot of ginger. The dual-RNAseq approach highlights responses of *P. myriotylum* that suggests features of a defensive strategy, and are perhaps another factor that may contribute to the variable success and durability of biological attempts to control diseases caused by *Pythium* spp.

Keywords: *Pythium oligandrum*, bioeffector, dual RNAseq, antagonism, ginger, *Pythium myriotylum*

INTRODUCTION

Biological control is a promising strategy to suppress diseases in an environmentally friendly manner whereby microorganisms, the biocontrol agents (BCAs), can antagonize a plant pathogen and prime plant defenses (Köhl et al., 2019). The application of biocontrol strategies in the field is often hampered by a lack of robustness where the previously observed disease protection effect is not maintained. One area needing further investigation is molecular aspects of the interaction between a plant pathogen and a biocontrol agent or in terms appropriate to their ecological interactions—the host or prey and the microbial antagonist.

One prominent biocontrol species is the oomycete *Pythium oligandrum,* which has several known mechanisms of antagonism contributing to its ability to control plant diseases as well as plant-mediated mechanisms. Recently, the review of Thambugala et al. (2020) summarized plant diseases caused by fungi or oomycetes that could be controlled by *P. oligandrum*. In particular, with regard to diseases caused by oomycetes, most of the diseases where *P. oligandrum* has a control effect are caused by *P. ultimum,* which causes diseases on a broad range of crop species (Thambugala et al., 2020). E.g., *P. oligandrum* has been reported to control pre- and post-emergence damping-off of wheat caused by *P. ultimum* (Abdelzaher et al., 1997). To our knowledge, there are no reports in the literature of antagonism of *P. oligandrum* toward *P. myriotylum* or to the control of diseases caused by *P. myriotylum*. An unpublished report showed that an isolate of *P. oligandrum* was antagonistic toward *P. myriotylum* in plate confrontation assays, but there was no control effect on Pythium soft-rot (PSR) of ginger caused by *Pythium myriotylum* (Le, 2016).

P. myriotylum is a broad host range pathogen that can infect a range of monocot and dicot plant species (van der Plaats-Niterink, 1981). One of these monocot hosts is ginger (*Zingiber officinale*) and recently, it was shown that *P. myriotylum* can be the most frequently found oomycete species from infected ginger rhizomes showing symptoms of PSR in the major ginger growing regions in China (Daly et al., 2021). Ginger is a spice and an important crop (De-Guzman and Siemonsma, 1999), with global production in 2019 estimated at approximately four million metric tons (FAOSTAT, 2019). Control strategies for PSR of ginger were comprehensively reviewed by Le et al. (2014) and are primarily based on chemical pesticide treatments. There is also potential for biological control methods such as a report of using a hypocrealean fungus *Trichoderma harzianum* to control PSR of ginger (Singh, 2011), but as reviewed by Le et al. (2014), the efficacies of the biological treatments appear lower in field conditions. In light of this, there is a need to identify novel biological control agents to control PSR of ginger.

Recently, the genome of the *P. oligandrum* ATCC 38472 strain was sequenced by long-read sequencing technology (Faure et al., 2020), and previously the *P. oligandrum* strains Po37 (Berger et al., 2016) and CBS 530.74 (Kushwaha et al., 2017) were sequenced by short-read sequencing technology. Note that the ATCC 38472 strain of *P. oligandrum* is used in the commercial oospore-containing product Polyversum®. In a comparative genomics analysis of *P. oligandrum* with other oomycetes, there was an expansion of carbohydrate-active enzyme (CAZy) families (Lombard et al., 2014) involved in the degradation of cellulose (Liang et al., 2020), perhaps related to the ability of *P. oligandrum* to degrade the cellulose in the cell wall of an oomycete host or prey. *P. oligandrum* is often referred to as a mycoparasite as its antagonism includes the coiling of *P. oligandrum* hyphae around hyphae of its host (Berry et al., 1993). *P. oligandrum* possesses an arsenal of hydrolytic enzymes such as cellulases and proteases that are important for antagonism toward particular hosts. The CAZyme content of *P. oligandrum* was described recently by Liang et al. (2020). The production of antimicrobial compounds by *P. oligandrum* also contributes to the antagonism (Benhamou et al., 1999). Interestingly, *P. oligandrum* can enter and colonize the root tissue before it rapidly degrades or degenerates itself (Benhamou et al., 2012). Presumably, this rapid degeneration of *P. oligandrum* can limit the damage that its cellulases can cause to the plant cell wall, of which cellulose is a major component. *P. oligandrum* appears to have a priming role in stimulating plant defense responses by microbe-associated molecular pattern (MAMP) production such as the elicitin-like protein oligandrin (Benhamou et al., 2001). Recently, type II NLP toxin proteins from *P. oligandrum* were found to suppress *Phytophthora capsici* infection (Yang et al., 2021). Expression of putative effectors was found when *P. oligandrum* interacted with heat-killed *Phytophthora infestans* mycelia (Horner et al., 2012), and these could potentially antagonize the host or mediate plant defense responses.

Dual transcriptomics simultaneously analyses the transcriptomes of two or more interacting organisms and is frequently used in host-pathogen interaction (Westermann et al., 2017). This can uncover expression differences in interacting organisms such as the upregulation of genes related to antagonism when three saprobic ascomycete fungi were cultured together (Daly et al., 2017). Dual-RNAseq was used to uncover the transcriptional responses of the

biocontrol rhizobacterium *Lysobacter capsici* and the pathogen *Phytophthora infestans* (Tomada et al., 2017). In this interaction, *L. capsici* genes involved in host colonization and degradation, and detoxification were upregulated while in contrast the host *P. infestans* was not considered to have mounted a major defensive or counter-attacking response but a transcriptomic response perhaps related to cell death (Tomada et al., 2017).

Here we investigated if a new isolate of *P. oligandrum* could antagonize *P. myriotylum* and used dual-RNAseq to measure the changes in gene expression that occurred when the two species confronted each other, alongside investigating if *P. oligandrum* had a control effect on the disease caused by *P. myriotylum*.

MATERIALS AND METHODS

Pythium Strains and Routine Culturing

P. oligandrum GAQ1 CGMCC No. 17470 (Wei et al., 2019) was isolated from soil from a field where infected ginger was growing in Laiwu district, Jinan City, Shandong Province, China. The molecular marker sequences of the GAQ1 isolate (MK774755.1, MZ891585.1, MZ891586.1 and MZ869812.1) have been deposited in GenBank. The *P. myriotylum* SWQ7 CGMCC No. 21459 isolate was described previously by Daly et al. (2021). The *Pythium* species were routinely cultured on 10% V8 juice agar at 25°C and maintained on V8 juice agar slants at 12°C.

Identification of *P. oligandrum* GAQ1

The CTAB protocol of Doyle (1990) was used to extract DNA from mycelial mats growing on V8 juice agar. The ITS region including the 5.8S rRNA subunit was amplified using the two universal primers ITS1 and ITS4 (White et al., 1990). The *CoxI* gene was amplified using OomCoxILevup and OomCoxI-Levlo primers (Robideau et al., 2011), the *CoxII* gene using FM66 and F58 primers (Martin and Tooley, 2003), and the β-tubulin gene using TUBUF2 and TUBUR1 primers (Kroon et al., 2004). PCR reactions were performed using a proof-reading polymerase with standard reaction and cycling conditions. All PCR products were run on an agarose gel to confirm the presence of an amplified product. The sequencing of the PCR products, BLAST searches and phylogenetic analysis were performed as described previously (Daly et al., 2021).

Visual Analysis of *P. oligandrum* and *P. myriotylum* Interaction

From fresh pre-culture plates on V8 juice media, an agar plug from the hyphal front was used to inoculate onto opposite ends of V8 agar on 9 cm plates and incubated at 25°C. For the macroscopic observation of the colonies, 20 mL of V8 agar was added to the plate, and for the microscopic visualization, 4.5 ml of V8 agar was added to the plates to provide a thinner layer. 5 mL of a 0.01% w/v trypan blue solution was used to cover the mycelium and agar surface. Trypan blue is a type of vitality stain, and it stains non-viable tissues with a blue color. The plates were gently rotated to ensure that the dye and mycelium were in full contact, and then incubated without shaking at room temperature for 3 min. Then the staining solution was decanted, and any remaining dye was pipetted off. The stained mycelial cultures were washed twice with 5 ml of water by incubating for 5 min without shaking. The staining was done after contact of the hyphal fronts and at various stages afterward. Microscopic visualization was performed using a brightfield microscope (NE900, Nexcope). At least three replicate cultures were used for the visualization work. A cryo-scanning electron microscope (cryo-SEM) (Quorum PP3010T system integrated onto a Hitachi SU8010 FE-SEM) was also used to visualize the interaction between *P. oligandrum* and *P. myriotylum*.

Dual-RNAseq Transcriptomics Experiment

Pre-cultures were grown on 10% V8 juice agar at 25°C. From the edge of the colony of *P. myriotylum* SWQ7 or *P. oligandrum* GAQ1, 0.5 cm mycelial plugs were transferred to 10% V8 juice agar covered with a 14.2 cm diameter polycarbonate membrane with a 0.1 μm pore size (GVS filter technology, Cat. no. 1215304) in a 14 cm Petri dish. The mycelial plugs were placed equidistant from each other and the edge of the Petri dish. The Petri dishes were inoculated with either two plugs of the same species (self-confrontation) or one plug from each species (mixed species confrontation). The cultures were incubated at 25°C in the dark, and mycelia from the confrontation zone was sampled as shown by the dotted-lines (**Supplementary Figure 1**). The mycelium was flash-frozen in liquid nitrogen before being ground using a Tissue Lyser. The sample was mixed with 1 mL of Trizol before adding 200 μl chloroform: isoamyl alcohol (24:1). The RNA was precipitated by mixing 0.8 volume of isopropanol with the aqueous phase from the previous step and the resulting pellet was washed once with 70% EtOH before resuspension in DEPC-treated water. The RNA was then cleaned up using a Tiangen Plant RNA extraction kit which included an on-column DNase treatment. RNA integrity was visualized using an agarose gel and a Bioanalyzer, and RNA quantity and purity using a NanoDrop spectrophotometer.

RNA Sequencing and Data Analysis

A paired library 2 × 150 bp was prepared using the TruSeq Stranded mRNA LT Sample Prep Kit (Illumina) and sequenced using the MGISEQ-2000 platform (MGI Tech Co. Ltd). The raw reads were processed using Trimmomatic (Bolger et al., 2014). The reads containing poly-N and the low-quality reads were removed to obtain the clean reads. Downstream analysis of the cleaned reads was performed using the following software tools in UseGalaxy.eu (Afgan et al., 2018). For all samples, reads that passed the filters were mapped using HISAT2 Galaxy Version 2.1.0 (Kim et al., 2019) onto a composite reference genome that consisted of the genomes of both species. The composite genome was made by concatenating into a single file those Fasta files made of the sequences for the scaffolds from *P. oligandrum* ATCC 38472 GCA_005966545.1_ASM596654v1_genomic.fasta (Faure et al., 2020) and *P. myriotylum* SWQ7 JAAS_PmSWQ7_1_0_assembly.fsa (PRJNA692555) (manuscript in preparation). The use of this composite genome facilitated

the exclusion from subsequent analysis of reads that were not unique to a particular species as these reads would not be uniquely mapped in the composite genome. Uniquely mapped read counts (at the MAPQ10 threshold) for each gene were calculated using HTSeq-count Galaxy version 0.9.1[1] (Anders et al., 2015). Read counts were calculated for each species separately using a genome annotation file for each species that contained the known gene exon coordinates for the genes for that species. The genome annotation file used for *P. oligandrum* was GCA_005966545.1_ASM596654v1_genomic.gff (Faure et al., 2020). Principal component analysis (PCA) indicated if the biological replicates were sufficiently similar for subsequent statistical analysis. The read counts of genes values for each species were subsequently used for statistical analysis using DESeq2 Galaxy Version 2.11.40.6 (Love et al., 2014). Note that the Cook's distance cut-off was not used for outlier filtering. The criteria for a differentially expressed (DE) gene were an FPKM > 10 in one condition, DESeq p_{adj} value < 0.05 and a DESeq log2 FC > 1 or < -1. FPKM normalized read counts were defined originally by Mortazavi et al. (2008). In our study, an FPKM value for a gene is the number of uniquely mapped (to the composite genome) fragments per kilobase of gene model per million uniquely mapped (to the composite genome) fragments onto gene models from a species. The RNA-seq reads from this project were submitted to the GEO database (GEO accession GSE179387).

Gene Annotations of *P. myriotylum* and *P. oligandrum* Genes

For the *P. oligandrum* genes, the annotations of Faure et al. (2020) for the ATCC 38472 strain were used. The *P. myriotylum* genes were functionally annotated primarily using InterProScan (Jones et al., 2014), and the CAZymes were annotated using dbCAN2 (Zhang et al., 2018; manuscript in preparation). The annotations for the *P. oligandrum* and *P. myriotylum* genes are listed in **Supplementary Table 1**.

qPCR Validation of Expression Differences in Dual-RNAseq

The qPCR validation was performed using RNA samples extracted from triplicate repeat cultures using the same experimental conditions and extraction method as described for the RNAseq experiment. cDNA was synthesized using the EasyScript one step gDNA removal and cDNA synthesis kit (TransGen Biotech) using random primers and 2.5 μg of total RNA according to the manufacturer's instructions. The qPCR was performed with a LightCycler 96 Instrument (Roche Life Sciences) using the ChamQ SYBR qPCR Master Mix (VAZyme). The cycling conditions were 95°C for 3 min, followed by 40 cycles of 95°C for 10 s and 60°C for 30 s. A melt curve analysis step was included to confirm the specificity of the primer pairs used. A standard curve of fivefold serial dilutions of the gDNA from *P. myriotylum* SWQ7 and *P. oligandrum* GAQ1 was used to determine the primer efficiency. The data was analyzed using the LightCycler 96 software. To quantify the abundance of the target in the cDNA samples, the relative standard curve method was used, and the expression levels of the gene of interest were normalized to two reference genes. A tubulin gene (*Pm_g6466.t1*) and a gapdh gene (*Pm_17033.t1*) were used for *P. myriotylum,* and a succinate dehydrogenase gene (*Po_g2472.t1*) and a g6pdh gene (*Po_g6912.t1*) were used for *P. oligandrum*. A pair of primers that amplified from gDNA was used to confirm the efficiency of the gDNA removal treatments for the gDNA from each species. Primers were designed using Primer-BLAST software (Ye et al., 2012) at NCBI and the primers used are listed in **Supplementary Table 2**. The primer efficiency for the primer pairs used was at least 86% and generally $> 90\%$, as detailed in **Supplementary Table 2**. Melt curve and agarose gel analysis were used to confirm the specificity of the primer pairs along with sequencing of the PCR products. A confirmation that the primers designed for *P. myriotylum* were unable to amplify from *P. oligandrum* and *vice-versa* was also performed.

Ginger Pot-Trial With *P. myriotylum* and *P. oligandrum* to Determine Biocontrol Effect

P. myriotylum was cultured on 10% V8 juice agar for 1–2 day, and then twice-autoclaved wheat seeds (which had previously been soaked overnight in water) were plated onto the mycelial mat of *P. myriotylum* and incubated for a further 7 d or until the seeds were fully covered with mycelia. The wheat seeds from the *P. myriotylum* cultures (after scraping away the excess mycelium on the outside of the seed) were used to inoculate the vermiculite near the roots at a depth of 2–3 cm. *P. oligandrum* was inoculated using four mycelia-covered balls from liquid culture. The use of the wheat seed inoculum method for *P. myriotylum* was based on the sorghum seed inoculum method of Le et al. (2016). To produce the *P. oligandrum* inoculum, *P. oligandrum* was incubated on 10% V8 juice agar for 1 day at 25°C, and then 20 agar plugs of a diameter of 0.5 cm were used to inoculate 50 mL of 10% V8 juice medium in a 150 mL flask, then incubated for 3 d at 25°C and 100 rpm. A mycelium-based inoculum was used for *P. oligandrum* because oospore production has not been optimized to achieve a yield that can practically be used as an inoculum. Ginger plants ("Laiwu" variety) were derived from tissue culture and transplanted to autoclaved vermiculite in 100 mL pots with an inverted Petri dish lid placed underneath each pot (to hold the water and prevent cross-contamination of the pathogen), and grown in 16 h light and 8 h dark cycles at 25°C in a growth chamber. Periodically, the ginger plants were given a water-soluble fertilizer (NPK 20-20-20 + TE). Eighteen ginger plants (grown in individual pots) were used for each of the six test or control conditions (**Table 1**). After 7 days from deflasking of the ginger seedlings, the experiment was initiated by inoculating three conditions with *P.* oligandrum, and two of these *P. oligandrum*-inoculated conditions were later inoculated with *P. myriotylum* at either 7 or 14 days while the third *P. oligandrum*-inoculated condition was not inoculated with *P. myriotylum*. Two of the other conditions were inoculated only with *P. myriotylum*

[1]http://www-huber.embl.de/users/anders/HTSeq/doc/count.html

TABLE 1 | Experimental design for ginger pot-trial with *P. myriotylum* and *P. oligandrum* to determine biocontrol effect.

	Controls		Test of 7 days between bioeffector and pathogen inoculations		Test of 14 days between bioeffector and pathogen inoculations	
Condition =	***P. oligandrum* alone inoculation**	**No inoculation**	***P. oligandrum* and *P. myriotylum* inoculations**	***P. myriotylum* inoculation**	***P. oligandrum* and *P. myriotylum* inoculations**	***P. myriotylum* inoculation**
n (plants) =	**18**	**18**	**18**	**18**	**18**	**18**
Day 0	*P. oligandrum* inoculation	—	*P. oligandrum* inoculation	—	*P. oligandrum* inoculation	—
Day 1	—	—	—	—	—	—
Day 2	—	—	—	—	—	—
Day 3	—	—	—	—	—	—
Day 4	—	—	—	—	—	—
Day 5	—	—	—	—	—	—
Day 6	—	—	—	—	—	—
Day 7	—	—	*P. myriotylum* inoculations		—	—
Day 8	—	—	—	—	—	—
Day 9	—	—	—	—	—	—
Day 10	—	—	—	—	—	—
Day 11	—	—	—	—	—	—
Day 12	—	—	—	—	—	—
Day 13	—	—	—	—	—	—
Day 14	—	—	—	—	*P. myriotylum* inoculations	

at either 7 or 14 days, and the last condition was not inoculated with either of the two species.

Two different *P. myriotylum* inoculation time-points were used to show the biocontrol effect after different time periods between pathogen and bioeffector inoculation, and from ginger plants of different ages. Seven days after inoculation with *P. myriotylum*, the plants were scored for 5 days with the following scale: 0 = plants remain green and healthy; 1 = leaf sheath collar discolored and lower leaves turned yellow; 2 = plants alive, but shoots either totally yellow or dead; and 3 = all shoots dead (Stirling et al., 2009). The disease index (DI) was calculated using the equation $DI = \Sigma\ 18_{i=1}\ Xi/54$ where DI is the disease index rating from 0 (healthy) to 1 (dead), Xi is the disease rating of the ith replicate (from 1 to 18), and 54 is equal to the number of replicates multiplied by the highest rating scale of 3. The disease control rate was calculated by the following equation: ("*P. myriotylum* inoculated plants" - "*P. oligandrum* and *P. myriotylum* inoculated plants")$_{diseaseindex}$/*P. myriotylum*$_{diseaseindex}$ × 100%.

To demonstrate that *P. oligandrum* could be recovered from around the roots of plants where there was a biocontrol effect on the PSR, the vermiculite from around the roots was collected into water containing the following five antimicrobial compounds: 100 μg/mL ampicillin, 50 μg/mL nystatin, 10 μg/mL pentachloronitrobenzene, 50 μg/mL rifampicin and 100 μg/mL carbenicillin. Serial dilutions of the suspension were plated onto 10% V8 agar containing the same antibiotics as above, and isolated colonies were sub-cultured onto V8 agar for DNA extraction using a CTAB-based extraction method. For PCR from the gDNA samples, ITS1 and ITS4 primers (White et al., 1990) were used to identify the species present. The 25 μL PCR reactions consisted of a 2X Taq Master Mix (VAZyme), 0.4 μM of each primer and 2 μL of gDNA template or water. The PCR cycling conditions were 95°C for 3 min, followed by 35 cycles of 95°C for 15 s, 56°C for 15 s, 72°C for 1 min, and then 72°C for 5 min. After agarose gel analysis, PCR products from the reactions using the ITS1 and ITS4 primers were sequenced using Sanger sequencing (Sangon Biotech).

RESULTS AND DISCUSSION

The GAQ1 Isolate Belongs to *P. oligandrum*

The *P. oligandrum* isolate GAQ1 was recovered from soil from a field where infected ginger with symptoms of PSR disease was growing. The ITS region sequence of GAQ1 (MK774755.1) is 99.87% similar (788/789) to the strain of *P. oligandrum* CBS 382.34 (AY598618.2) used in the taxonomic monograph of the *Pythium* genus of van der Plaats-Niterink (1981). For three other molecular marker sequences compared to the CBS 382.34 strain, the GAQ1 isolate was 99.69% similar to the *CoxI* sequence, 99.81% similar to the *CoxII*, and 99.87% similar to the β-*tubulin* sequence. In a multi-locus phylogenetic analysis (using the ITS region, *CoxI*, *CoxII* and β-*tubulin* sequences), GAQ1 formed a clade with the *P. oligandrum* CBS 382.34 strain and with genome-sequenced *P. oligandrum* strains (**Supplementary Figure 2**). Previous reports of the ability of *P. oligandrum* to control other diseases caused by *Pythium* spp. such as *P. ultimum*, e.g., Abdelzaher et al. (1997), prompted an investigation into whether the newly recovered isolate of *P. oligandrum* could antagonize

P. myriotylum, and what transcriptional changes occurred during their interaction, and whether *P. oligandrum* could control PSR of ginger caused by *P. myriotylum*.

Visualization of *P. oligandrum* and *P. myriotylum* Interaction

We were interested in understanding if the *P. oligandrum* GAQ1 isolated from soil where infected ginger was growing could control diseases of ginger such as *Pythium* soft rot. Previously, we recovered a virulent isolate of *P. myriotylum* named SWQ7 that could infect and kill ginger plants (Daly et al., 2021). To understand whether the GAQ1 isolate could antagonize *P. myriotylum*, we used plate-based confrontation assays and microscopy coupled with vitality staining using trypan blue. Both *P. oligandrum* and *P. myriotylum* had similar growth rates on V8 medium. Before contact of the mycelial fronts of both species, blue staining with trypan blue (indicating the presence of non-viable mycelium) was not observed from either species. Several hours after contact, the confrontation zone showed staining with the trypan blue (**Figure 1**), indicating a loss of viability in some of the hyphae. As the incubation of other cultures were continued for longer periods, the blue-stained area was seen to expand to cover the entire side of the plate where *P. myriotylum* had previously been growing on. The pattern of the trypan blue staining suggested successful antagonism by *P. oligandrum* of *P. myriotylum* as the loss of viability of the hyphae only appeared to occur in the part of the cultures where *P. myriotylum* was growing before the confrontation with *P. oligandrum*. Of note, when *P. oligandrum* or *P. myriotylum* was confronted with itself, the occurrence of trypan blue staining was not observed (**Figure 1**).

The brightfield microscopy revealed that the trypan blue staining occurred on the thicker hyphae which are generally found from *P. myriotylum* while the thinner hyphae of *P. oligandrum* lack the trypan blue stain. **Figures 2A,B** demonstrate that the size of *P. myriotylum* hyphae is larger than those of *P. oligandrum*. Clear occurrences of hyphal coiling were observed where the thinner hyphae of *P. oligandrum* coiled around the thicker trypan blue-stained hyphae of *P. myriotylum*, with also occurrences of penetration of the *P. myriotylum* hyphae by *P. oligandrum* (**Figure 2C**). Cryo-SEM was used to obtain higher magnification and resolution images of the hyphal coiling (**Figures 2D–F**). In the brightfield images, there were also occurrences of the spiny oogonia on the part of the plate where *P. myriotylum* had previously been growing. Spiny oogonia are one of the characteristic features of *P. oligandrum* and related species (van der Plaats-Niterink, 1981) and are not formed by *P. myriotylum*. The observation here of the spiny oogonia further demonstrates the colonization by *P. oligandrum* of the part of the plate where *P. myriotylum* had previously been growing.

The analysis here suggests a clear parasitic interaction of the *P. oligandrum* GAQ1 isolate on *P. myriotylum*. The macroscopic observations of the loss of vitality of the *P. myriotylum* hyphae after contact with *P. oligandrum* coupled with the microscopic

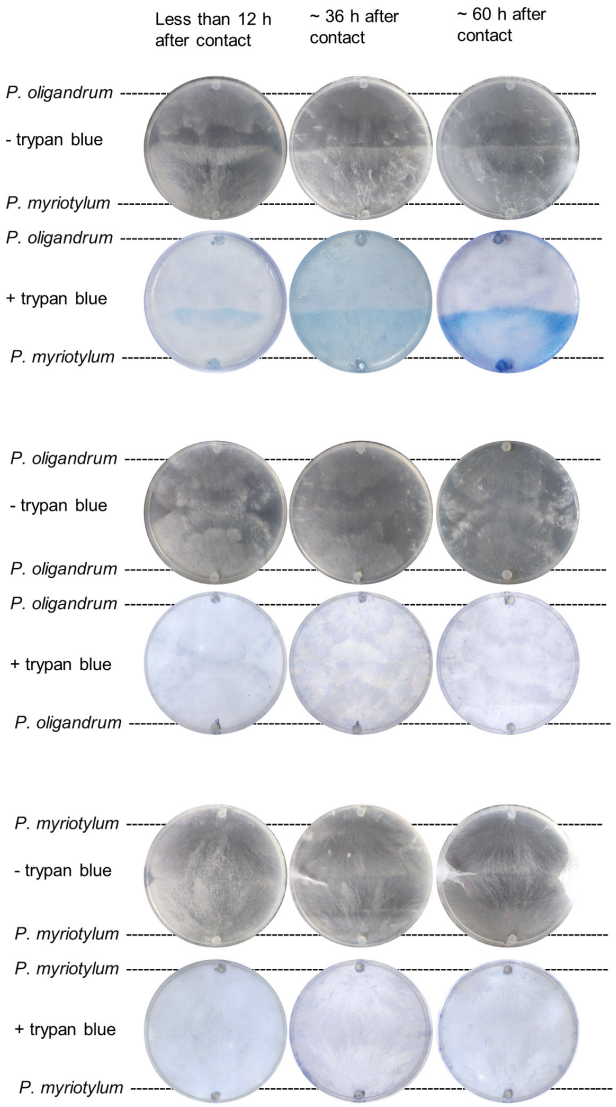


FIGURE 1 | Representative images of trypan blue staining of mycelia from *P. myriotylum* and *P. oligandrum* self- or mixed-species confrontation demonstrating that *P. myriotylum* mycelium becomes non-viable in the presence of *P. oligandrum*. Images before and after trypan blue staining of the cultures are shown from less than 12 h, from ~ 36 h and from ~60 h after contact of the hyphal fronts of the colonies. The brightness and contrast of the images was adjusted consistently. The images from the control self-confrontation cultures demonstrate that the blue-colored staining does not appear in the self-confrontation.

observations of hyphal coiling support this parasitic mechanism. Previously, there are reports of the *P. oligandrum* isolate CGH1 antagonizing multiple *Pythium* spp. with observances of hyphal coiling and hyphal penetration (Berry et al., 1993). Interestingly, in the study of Berry et al. (1993), a lack of antagonism from the *P. oligandrum* isolate toward the plant pathogen *P. aphanidermatum* was found, and instead *P. aphanidermatum* coiled around the hyphae of *P. oligandrum*. The two species of *P. aphanidermatum* and *P. myriotylum* are relatively closely

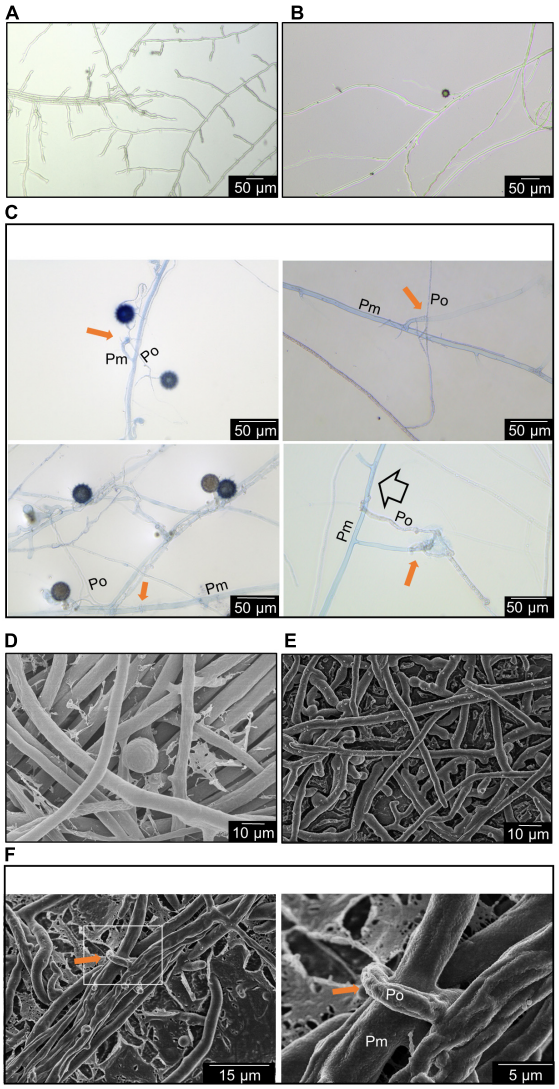


FIGURE 2 | Representative features of the interaction between *P. myriotylum* (Pm) *and P. oligandrum* (Po). Here hyphae are imaged in the region where the two colonies have recently made contact using either brightfield microscopy coupled with trypan blue staining **(A–C)** or cryo-scanning electron microscopy **(D–F)**. **(A)** *P. myriotylum* from self-confrontation, **(B)** *P. oligandrum* from self-confrontation with spiny oogonia visible, **(C)** coiling of *P. oligandrum* hypha around the hypha of *P. myriotylum* and apparent penetration by *P. oligandrum* of *P. myriotylum* hypha. The brightfield images of the cultures from where *P. oligandrum* and *P. myriotylum* confront each other were stained with trypan blue. **(D)** *P. myriotylum* from self-confrontation, **(E)** *P. oligandrum* from self-confrontation, **(F)** coiling of a *P. oligandrum* hypha around a hypha of *P. myriotylum* at two different magnifications. The cryo-SEM images were taken from cultures growing on polycarbonate membranes using the same experimental set-up as was used for the Dual-RNAseq (see **Supplementary Figure 1** for images of these cultures). The solid amber arrow indicates examples of hyphal coiling and the empty arrow indicates an example of penetration.

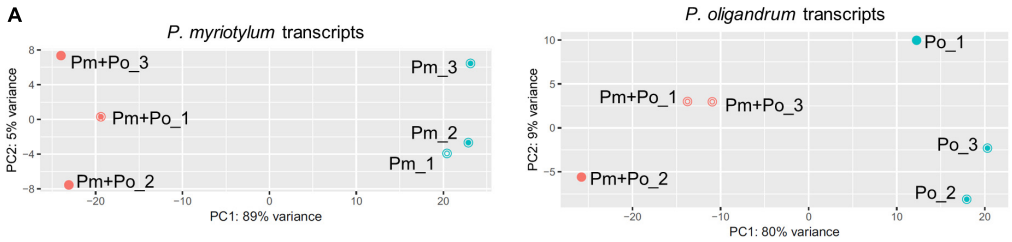
related, being from the *Pythium* clades A and B, respectively (Lévesque and De Cock, 2004), but in our study, we did not find occurrences of *P. myriotylum* hyphae coiling around the hyphae of the *P. oligandrum* GAQ1 isolate.

Overview of Transcriptome Analysis of the Interaction Between *P. oligandrum* and *P. myriotylum*

The interaction between *P. oligandrum* and *P. myriotylum* as observed macro- and microscopically was used as the basis to sample mycelia for transcriptomic analysis when *P. myriotylum* and *P. oligandrum* confronted each other or themselves. The part of the mycelia indicated by the black dotted-line in **Supplementary Figure 1**. were sampled for RNAseq which corresponds to an early time-point in the interaction between *P. oligandrum* and *P. myriotylum*. The *P. oligandrum* GAQ1 isolate had a flatter colony morphology with a notable chrysanthemal pattern to the mycelia, compared to the *P. myriotylum* SWQ7 isolate, which had a fluffier colony appearance with considerable numbers of aerial hyphae (**Supplementary Figure 1**). Approximately 20 million reads were obtained from each sample with ∼ 90% of the reads aligning to one location in a concatenation of the genome assemblies of the two species (**Supplementary Table 3**). When the reads that aligned uniquely to gene models from each species from the mixed species confrontation samples were counted, similar numbers of reads were counted on gene models for each species, with the biggest difference being only twofold in replicate two (**Supplementary Table 3**). From the single-species cultures of *P. oligandrum* or *P. myriotylum*, very few reads were counted on the gene models from the species not found in the culture, i.e., at between 0.005 and 0.06% of the reads that were counted on the gene models of the species present in the single culture (**Supplementary Table 3**). The above analysis indicated that the genes from the two species are sufficiently different in sequence to be distinguished from each other. The clustering pattern from a principal component analysis of the expression levels of the genes from each species showed separate clustering of the replicates from the self-confrontation compared to when the two species confronted each other (**Figure 3A**). In terms of the number of differentially expressed genes, there were 707 genes upregulated and 315 genes downregulated in *P. myriotylum* in the interaction with *P. oligandrum*. In *P. oligandrum*, there were 433 genes upregulated and 122 genes downregulated in the interaction with *P. myriotylum* (**Figure 3B** and **Supplementary Table 1**).

Transcriptomic Aspects of Parasitic Mechanisms From *P. oligandrum*

Cellulose is a key component of the oomycete cell wall (Mélida et al., 2013), and a potential major target of any antagonist of an oomycete like *P. myriotylum*. Based on CAZy, signal peptide and transmembrane domain annotations, *P. oligandrum* has 50 putative host/prey targeted cellulases. Host/prey targeted cellulases were considered those cellulases which contained a signal peptide and lacked transmembrane domains indicating that they are likely secreted into the extracellular environment (where they could degrade cellulose found in the cell walls of both a plant and oomycete host) and are unlikely to be involved in remodeling of the *P. oligandrum* cell wall. These host/prey targeted cellulases were from CAZy families GH1, GH3, GH5, GH6, GH7, GH30_1, GH131 and AA9. Of the putative host/prey



B

Gene ontology (GO) term	Pm ↑ (707 genes)	Pm ↓ (303 genes)	Po ↑ (433 genes)	Po ↓ (122 genes)
GO:0006508 – proteolysis			■	
GO:0008236 – serine-type peptidase activity			■	
GO:0004553 – hydrolase activity, hydrolyzing O-glycosyl compounds		■	■	
GO:0004601 – peroxidase activity			■	
GO:0016491 – oxidoreductase activity		■	■	
GO:0006303 – double-strand break repair via nonhomologous end joining	■			

FIGURE 3 | (A) Overview of the clustering pattern from principal component (PC) analysis of the transcripts from the replicate cultures from *P. oligandrum* (Po) and *P. myriotylum* (Pm) self- or mixed-species confrontations, **(B)** summary of interesting gene ontology (GO) terms enriched in the differentially expressed genes in *P. myriotylum* and *P. oligandrum* when they confront each other compared to a self-confrontation. The full list of enriched GO terms can be found in **Supplementary Table 4**.

targeted cellulases, 11/50 were significantly upregulated in the interaction with *P. myriotylum,* and notably, none of these 50 was significantly downregulated (**Supplementary Table 1C**). In general, the *P. oligandrum* putative host/prey targeted cellulases were expressed but were not upregulated to a high level in the presence of *P. myriotylum*. Only two of the upregulated host/prey targeted cellulases had an FPKM > 100 (two GH5_14 members that likely encode for β-glucosidase activity). Two members of the GH30_1 sub-family that likely also encode for β-glucosidase were constitutively expressed with FPKM > 100 but there were no good examples of upregulated host/prey targeted cellulases that likely have activity on the cellulose polymer (e.g., cellobiohydrolases or endoglucanases). *P. oligandrum* protease or peptidase encoding genes were a major category of genes that were upregulated in the interaction with *P. myriotylum* e.g., the GO term for—proteolysis (GO:0006508) was enriched in the *P. oligandrum* genes upregulated in the interaction with *P. myriotylum* (**Figure 3B**). The putative role of *P. oligandrum* proteases will be described in more detail in a later section.

The limited upregulation of the host/prey targeted cellulases in *P. oligandrum* in the confrontation with *P. myriotylum* contrasts somewhat with their expression in *P. oligandrum* in the interaction with *P. infestans* from the study of Liang et al. (2020). Here the majority of the AA9 cellulases were upregulated (with a subset expressed at a high level) at 12 and 24 h after contact between *P. oligandrum* and *P. infestans* compared to control samples before contact was made (Liang et al., 2020). Differences in the timing of the sampling for expression analysis, using a different *P. oligandrum* strain (CBS530.74), or other hosts (e.g., differences in the cell wall composition between the hosts) could contribute to the differences in cellulase gene expression when *P. oligandrum* confronted *P. myriotylum* compared to when *P. infestans* was confronted. The cellulolytic and protease gene expression measured from the GAQ1 isolate likely contributes to the effects seen in the microscopy images whereby there are clear signs of a loss of viability of the *P. myriotylum* mycelium from the trypan blue staining (**Figures 1**, **2**) and examples of hyphal penetration by *P. oligandrum* of the *P. myriotylum* hyphae (**Figure 2C**).

Apart from genes involved in a parasitic mechanism, several known genes from *P. oligandrum* with a role in mediating the biocontrol of diseases via stimulation of plant defenses, were highly expressed at a constitutive level. e.g., oligandrin *Po_g9166.t1* (*Oli−D1*) and *Po_g9168.t1* (*POD-2*) (there is no evidence that the presence of *P. myriotylum* downregulated these genes). Oli-D1 and POD-2 are elicitin/elicitin-like proteins, and seven other *P. oligandrum* elicitin-like genes were upregulated in the confrontation with *P. myriotylum,* but their expression level was 10–100-fold less than *Oli-D1* and *POD-2* (**Supplementary Table 1**).

Features of the Transcriptomic Response of *P. myriotylum* When Parasitized by *P. oligandrum*

A subset of *P. myriotylum* putative cellulase genes were upregulated in the confrontation with *P. oligandrum*. In contrast, to the *P. oligandrum* upregulated cellulases described in the previous section, a smaller number of *P. myriotylum* cellulases were upregulated, and at least half seemed likely to have a cellulose remodeling or intracellular role rather than as part of a counter-attacking defensive strategy. Six *P. myriotylum* genes annotated as putative cellulases were upregulated: *Pm_g1192.t1* (GH5_14), *Pm_g1197.t1* (GH5_14),

Pm_g7072.t1 (GH5_20), *Pm_g7136.t1* (GH5_20), *Pm_g7449.t1* (AA9), and *Pm_g14120.t1* (GH5_20). Two of the putative cellulases were annotated with a signal peptide (*Pm_g1192.t1* and *Pm_g7449.t1*) and three of those without a signal peptide were annotated with transmembrane domains (*Pm_g7072.t1*, *Pm_g7136.t1,* and *Pm_g14120.t1*), suggesting these three are membrane-bound proteins. The six upregulated *P. myriotylum* putative cellulases were annotated as either GH5_14, GH5_20, or AA9 family members. The characterized activities within the GH5_14 family include β-glucosidase and exo-β-1,3-glucosidase activities.[2] The β-glucosidase activity could act on the cellobiose released from cellulose but cellulose does not contain β-1,3 linked glucans. The GH5_20 members appear to only be found in Stramenopiles and appear most closely related to GH5_1 members, of which many have been characterized as endo-β-1,4-glucanases (Aspeborg et al., 2012) but no GH5_20 members appear to have been characterized to date. The expression levels of the three putative cellulases from the GH5_20 family encoded by *Pm_g7072.t1*, *Pm_g7136.t1,* and *Pm_g14120.t1* were validated by qPCR (**Figure 4A**). For the other CAZy family found amongst the upregulated cellulases; generally, proteins containing an AA9 domain have oxidative activity toward crystalline cellulose (Quinlan et al., 2011). The putative transmembrane domain-containing cellulases upregulated in *P. myriotylum* when confronted with *P. oligandrum* may remodel the cellulose in the *P. myriotylum* cell wall as part of a defensive response to the *P. oligandrum* antagonism. In mycoparasitic interactions involving *Trichoderma atroviride*, the remodeling of prominent fungal cell wall polysaccharides chitin and chitosan was important for the mycoparasitic ability of the fungus by contributing to self-defense (Kappel et al., 2020).

A subset of nine *P. myriotylum* elicitin-like putative pathogen-associated molecular pathogen (PAMP) molecules were upregulated in the confrontation with *P. oligandrum*. The upregulation of the elicitin-like *Pm_g6122.t1* was validated by qPCR (**Figure 4A**). Of note, other elicitin-like proteins are highly constitutively expressed, such as *Pm_g10365.t1* with an FPKM ∼ 20,000 that are more likely to function in *P. myriotylum* in the major primary role of the elicitin-like proteins in scavenging for sterols (Janků et al., 2020). If the upregulation of a subset of the *P. myriotylum* elicitin-likes found in the *in vitro* confrontation was replicated in the tri-partite interaction of the two *Pythium* spp. and a plant host of *P. myriotylum* (e.g., ginger), there is the potential to modulate plant immunity and the virulence of *P. myriotylum*, as elicitin-like proteins from other oomycete species can function as PAMPs (Derevnina et al., 2016). This, of course, would be an unintentional consequence of the interaction of *P. oligandrum* on *P. myriotylum* as a bioeffector does not intentionally alter the virulence of the plant pathogen it antagonizes as instead, the plant pathogen is a nutrient source or a competitor for ecological resources.

The GO term for double-strand break repair via non-homologous end-joining (GO:0006303) was enriched in the genes upregulated in *P. myriotylum* in the interaction with *P. oligandrum* (**Supplementary Table 4**). The two genes responsible for this enrichment, *Pm_g12183.t1* and *Pm_g10950.t1*, were annotated with domains found in Ku70/Ku80 proteins (PF02735, PF03730 and PF03731). Ku proteins are evolutionarily conserved proteins that mediate repair of double-stranded DNA by non-homologous end-joining of the broken ends of the DNA (Doherty and Jackson, 2001). Although these two genes were also expressed in *P. myriotylum* when it confronts itself, their upregulation in the confrontation with *P. oligandrum* suggests there could be a greater requirement to repair double-strand breaks.

In *P. myriotylum*, there were relatively few protease-encoding genes upregulated in the confrontation with *P. oligandrum*. Instead, one stand-out part of the interaction between *P. myriotylum* and *P. oligandrum* was the high upregulation of *P. myriotylum* putative protease inhibitors along with the upregulation or constitutive expression of *P. oligandrum* proteases, which will be described in detail in the next section.

High Upregulation of a Sub-Set of *P. myriotylum* Kazal-Type Protease Inhibitors and *P. oligandrum* Peptidases

In the *P. myriotylum* genome, there are 23 genes annotated with a Kazal-type protease inhibitor domain (**Supplementary Table 1** and **Figure 5A**). One of the most striking features of the genes induced in the interaction between *P. myriotylum* and *P. oligandrum* was the upregulation of six of the *P. myriotylum* genes annotated with a Kazal-type protease inhibitor domain in the interaction with *P. oligandrum* (*Pm_g8255.t1*, *Pm_g8256.t1*, *Pm_g8257.t1*, *Pm_g8263.t1*, *Pm_g8265.t1,* and *Pm_g8266.t1*). Three of these genes (*Pm_g8256.t1*, *Pm_g8257.t1,* and *Pm_g8265.t1*) were induced greater than 1,000-fold to a high FPKM value of greater than 1000 FPKM. All six of these Kazal-type protease inhibitor genes were annotated with a signal peptide, and only *Pm_g8257.t1* is predicted to possesses a transmembrane domain, indicating that the other five are likely secreted into the environment around *P. myriotylum*. The induction of the genes annotated with a Kazal-type protease inhibitor domain was validated by qPCR (**Figure 5B**). Of the other Kazal-type protease inhibitor genes that were not upregulated, one of the genes (*Pm_g16340.t1*) appeared to be constitutively expressed when *P. myriotylum* was confronted by itself or by *P. oligandrum*, and two other Kazal-type protease inhibitor genes (*Pm_g8268.t1* and *Pm_g1759.t1*) were downregulated when *P. myriotylum* was confronted by *P. oligandrum*. Overall, for the *P. myriotylum* genes annotated with a Kazal-type protease inhibitor domain, the dominant feature was high upregulation of a subset of these when *P. myriotylum* was confronted by *P. oligandrum*. Of the genes upregulated in *P. myriotylum* in the interaction with *P. oligandrum*, the largest fold-change upregulated and highest expression was seen for these genes annotated as Kazal-type protease inhibitors.

In contrast to the *P. myriotylum* Kazal-type protease inhibitor genes, in *P. oligandrum*, none of the 13 genes annotated with a Kazal domain were differentially expressed, while one gene annotated with a Kazal domain (*Po_g7532.t1*) appeared

[2]http://www.cazy.org/GH5_14_characterized.html

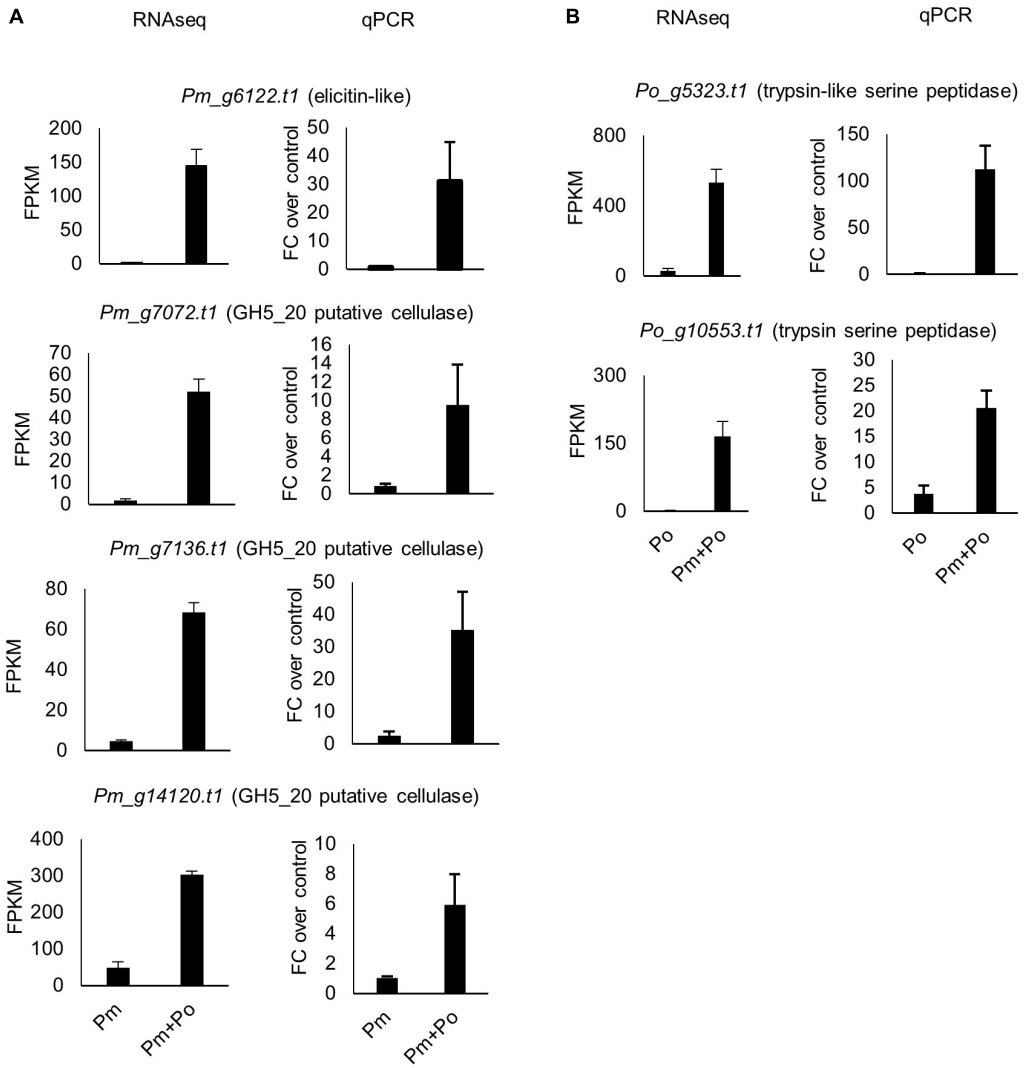


FIGURE 4 | Expression levels from RNAseq and qPCR for selected **(A)** *P. myriotylum* genes and **(B)** *P. oligandrum* genes upregulated in the mixed-species confrontation compared to the self-confrontation. The error bars represent standard errors ($n = 3$). The expression levels of *P. myriotylum* were normalized using the two reference genes *Pm_g6466.t1* (tubulin) and *Pm_17033.t1* (gapdh) and the expression levels of *P. oligandrum* were normalized using *Po_g2472.t1* (succinate dehydrogenase) and *Po_g6912.t1* (g6pdh). FC, fold change; Pm, *P. myriotylum*; Po, *P. oligandrum*.

to be constitutively expressed to a relatively high level of ∼500 FPKM. This gene, *Po_g7532.t1*, is a match to two EST sequences (EV245133.1 and EV243682.1) from an EST library sequenced from an *in vitro* interaction between *P. oligandrum* with heat-killed *P. infestans* mycelia (Horner et al., 2012). Another EST from the same library EV245779.1, annotated with a Kazal domain, matched *Po_g10361.t1*, which was expressed constitutively in the conditions at ∼70 FPKM.

Kazal-type protease inhibitors can inhibit serine proteases (Tian and Kamoun, 2005) and whether there was any upregulation of the class of proteases in *P. oligandrum* that the *P. myriotylum* Kazal-type protease inhibitors could inhibit was of interest. Interestingly, the GO enrichment analysis showed that the GO term for serine-type peptidase activity (GO:0008236) was enriched in the *P. oligandrum* genes upregulated in the confrontation with *P. myriotylum* (**Figure 3B**). The expression level of *P. oligandrum* putative serine proteases was assessed using the following Pfam annotations: Pfam00089 (Trypsin), Pfam13365 (Trypsin-like peptidase domain), Pfam00082 (Subtilase family), Pfam00450 (Serine carboxypeptidase), and Pfam05577 (Serine carboxypeptidase S28). Of these 142 *P. oligandrum* serine proteases, there were 23 upregulated when *P. oligandrum* confronted *P. myriotylum* compared to when *P. oligandrum* confronted itself. All of these 23 putative serine proteases were annotated with a signal peptide sequence (**Supplementary Table 1**). The upregulation of the peptidase with one of the largest fold changes in upregulation (*Po_g10553.t1*, a trypsin family serine protease) and the peptidase that had the highest expression level in the confrontation with *P. myriotylum*

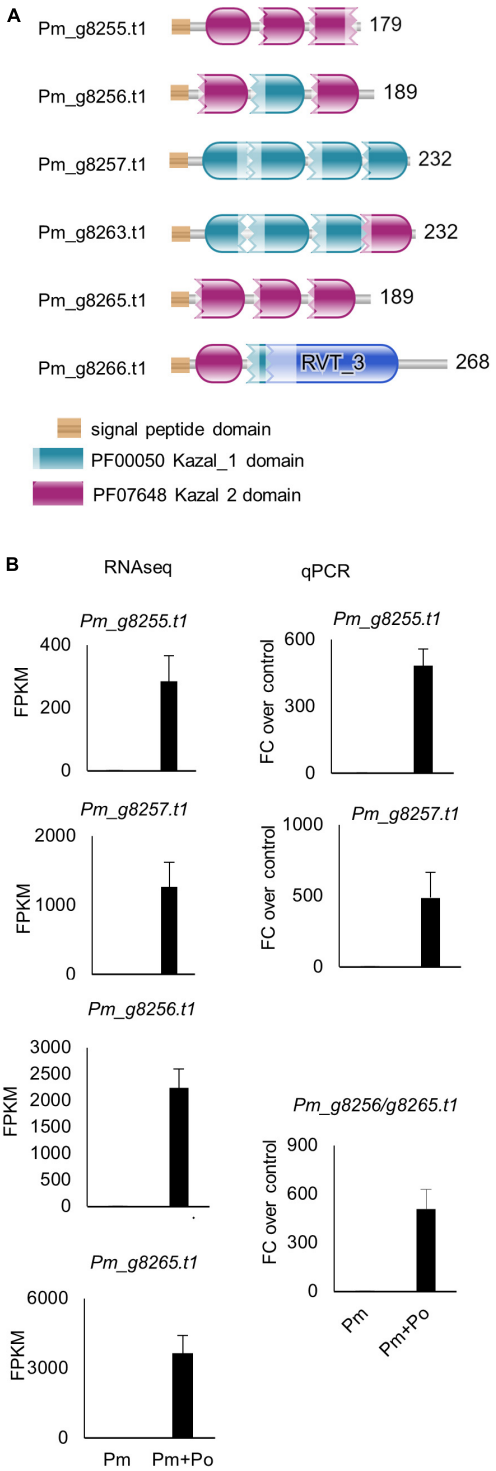


FIGURE 5 | (A) Protein sequences of the confrontation-upregulated *P. myriotylum* Kazal-type protease inhibitors highlighting the locations of the Kazal domains. **(B)** Expression levels from RNAseq and qPCR of a sub-set of the *P. myriotylum* Kazal-type protease inhibitors upregulated in the confrontation with *P. oligandrum* compared to the self-confrontation. The error bars represent standard errors ($n = 3$). The expression levels of *P. myriotylum* were normalized using two reference genes *Pm_g6466.t1* (tubulin) and *Pm_17033.t1* (gapdh). Pm, *P. myriotylum*; Po, *P. oligandrum*.

(*Po_g5323.t1*, a trypsin-like serine protease) was validated by qPCR (**Figure 4B**).

There is potential for these proteins with Kazal-type protease inhibitor domains to be involved in inhibition of the serine peptidases upregulated or constitutively expressed in *P. oligandrum*. Previously in the interaction between the biocontrol rhizobacterium *L. capsici* and the pathogen *P. infestans* (Tomada et al., 2017) there was an upregulation of a subset of Kazal-type protease inhibitors, and it was hypothesized that these could be involved in inhibiting lytic activities from *L. capsici*. Although, notably, Tomada et al. (2017) did not consider overall an active response from *P. infestans* to the antagonism from *L. capsici*. Consistent with a prominent role for *P. oligandrum* proteases in the antagonism of *P. myriotylum* was a lack of upregulation of the *P. oligandrum* Kazal-type protease inhibitors, as their upregulation could inhibit the *P. oligandrum* protease activity. Horner et al. (2012) considered the ESTs annotated with a Kazal domain as part of a defense or counter-defensive response of *P. oligandrum* to heat-killed *P. infestans* mycelia. However, in the confrontation between *P. oligandrum* and the living *P. myriotylum* mycelia, the *P. oligandrum* putative Kazal-type inhibitors appear to play a minor role.

While the transcriptomic response of *P. myriotylum* in the confrontation with *P. oligandrum* suggested aspects of a defensive or counter-attacking strategy, this appears insufficient to defend against *P. oligandrum* as clearly in the plate confrontation assays, *P. oligandrum* caused progressive loss of viability in the hyphae of *P. myriotylum* (**Figure 1**). As the ability to antagonize on plate cultures suggested it could have a control effect on diseases caused by *P. myriotylum*, the ability to control the Pythium soft-rot of ginger caused by *P. myriotylum* was investigated using a pot-trial experiment.

P. oligandrum GAQ1 Has Control Effect on Pythium Soft-Rot of *P. myriotylum* Infected Ginger

As *P. oligandrum* demonstrated an ability to antagonize *P. myriotylum* in plate confrontation assays, its ability to control PSR of ginger was investigated. *P. oligandrum* was inoculated first followed by inoculations of *P. myriotylum* at two time points after the inoculation of *P. oligandrum*, and this also led to inoculations of *P. myriotylum* onto plants of two different ages and sizes. Clearly from the data on the above-ground disease symptoms in **Figure 6**, the plants inoculated with *P. myriotylum* developed the disease symptoms. The disease was more severe in the younger and smaller ginger plants as a higher proportion of the plants developed the more severe symptoms whereby all the aerial parts of the plant appeared dead (**Figure 6A**). In both of the time-points analyzed, the presence of *P. oligandrum* resulted in a clear and significant ($P < 0.001$) reduction in the disease symptoms whereby > 80% of the plants had no disease symptoms (**Figure 6A**). The disease control rate was ∼80% in both inoculation time points (**Figure 6B**). To show that *P. oligandrum* could be recovered from plants where there was a biocontrol effect on the PSR, *P. oligandrum* was re-isolated

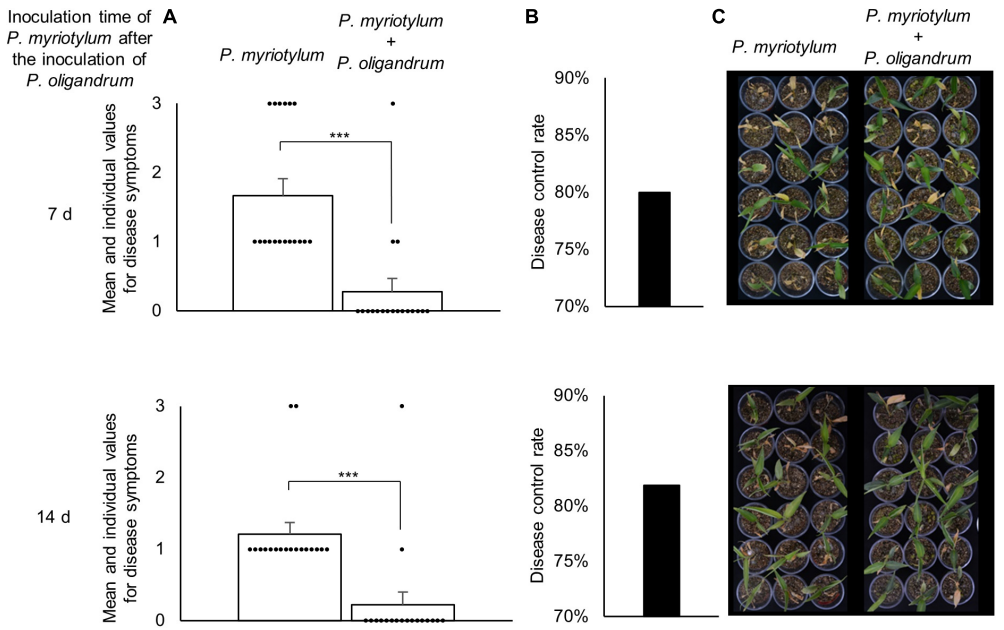


FIGURE 6 | Biocontrol effect of *P. oligandrum* on Pythium soft-rot (PSR) of ginger disease caused by *P. myriotylum* from two different inoculation time-points of *P. myriotylum* after the inoculation of *P. oligandrum*. **(A)** The mean values and raw data from the scoring of the disease symptoms (scale from 0 to 3) in the 18 replicate plants for each condition. ***corresponds to a $P \leq 0.001$ from a Mann-Whitney Rank Sum Test ($n = 18$). **(B)** The disease control rate of *P. oligandrum* on PSR of ginger. **(C)** Images of the 18 replicate plants inoculated with *P. myriotylum or P. myriotylum* and *P. oligandrum*. The data and images were recorded 11 d after the inoculation of the plants with *P. myriotylum*. The scoring for the disease symptoms was always zero for the plants that were inoculated with only *P. oligandrum,* and for the plants that were not inoculated with either of the *Pythium* spp.

from around the roots that had been infected with *P. myriotylum* (**Supplementary Figure 3**).

A similar experiment using a different *P. oligandrum* isolate to control PSR soft-rot of ginger caused by a *P. myriotylum* isolated from infected ginger in Australia did not reduce the disease symptoms on ginger plants (Le, 2016). It was speculated by Le (2016) that the lack of a control effect may have been in part due to poor colonization of the soil by the *P. oligandrum* isolate (vermiculite was used to grow the ginger plants inoculated with GAQ1) or the ineffective application of the oospore suspension by soil drenching (mycelia of the GAQ1 isolate was added directly into the vermiculite). However, there is support for *P. oligandrum* to control other diseases caused by *P. myriotylum* where previously it was shown that *P. oligandrum* was able to reduce soybean damping-off caused by *P. myriotylum* (You et al., 2019).

Part of the rationale for the suite of analyses undertaken in this study involving microscopy, dual-transcriptomics, and biocontrol was to look for further insights into why biocontrol strategies in the field are often hampered by a lack of robustness or durability. Of course, dominant reasons for this include interactions between the bioeffector and a complex microbiome, varying abilities of the bioeffector to propagate, and changes in host strain or isolate (i.e., the plant pathogen) from year to year. The biocontrol effect from the controlled laboratory conditions is promising, with a disease control rate of ∼80% for the application of *P. oligandrum* to control PSR of ginger, although the ability of *P. oligandrum* to suppress PSR of ginger in field conditions has yet to be tested. Suppose one were to extrapolate the microscopic and transcriptomic observations from the *in vitro* confrontation assays to the tri-partite interaction in the pot trial experiment. In that case, there is likely *P. oligandrum* parasitism or predation of *P. myriotylum* contributing to the suppression of the disease. The defensive or counter-attacking strategies suggested from the dual-transcriptomics, such as the Kazal-type protease inhibitor upregulation, do not appear to manifest themselves into a successful defense. In the *in vitro* plate confrontation assays, *P. oligandrum* progressively parasitized *P. myriotylum* mycelia, and a relatively high control rate of ∼80% was maintained in the pot-trial experiment. However, these defensive or counter-attacking strategies suggested by the dual-transcriptomics are another factor that is worthwhile in considering for the robustness and durability of biocontrol strategies in the field.

CONCLUSION

The *P. oligandrum* isolate GAQ1 is a promising bioeffector of PSR disease of ginger caused by *P. myriotylum* and has parasitic abilities that likely contribute to the control of PSR disease of ginger. Likely side-by-side comparisons of different *P. oligandrum* strains would be required to determine whether the GAQ1 isolate is superior for control of particular diseases (e.g., whether *P. oligandrum* GAQ1 which was isolated from a field with infested ginger is superior to other strains in controlling the PSR of ginger). For field-based application, an

oospore-based preparation of *P. oligandrum* GAQ1 would likely ensure better shelf-life of the bioeffector compared to the use of a mycelial-based inoculum. The use of dual-transcriptomics is a powerful approach to uncover differentially expressed genes in the interaction between an antagonist and its prey or host. The differentially expressed genes such as Kazal-type protease inhibitors are promising candidates to investigate further molecular aspects of the antagonism and defensive responses between host and prey.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

PD, SC, QZ, JZ, JL, and TX performed the experiments. PD, DF, and JM performed bioinformatic analyses. PD and SC wrote the manuscript and analyzed the results. XW and XS sampled the soil for microbe isolation. LW obtained funding for the research. TS, DZ, SD, MJ, ID, and LW contributed to the discussion and manuscript revision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported by the National Natural Science Foundation of China (32050410305), the China Agriculture Research System of MOF and MARA (CARS-24-C-01), the Jiangsu Agricultural Science and Technology Innovation Fund [CX(18)2005], and the China Postdoctoral Science Foundation (2020M671388).

ACKNOWLEDGMENTS

Assistance from OE Biotech company for next-generation sequencing and the electron microscopy facility staff at Nanjing Agricultural University for the microscopy work is acknowledged.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.765872/full#supplementary-material

REFERENCES

Abdelzaher, H. M. A., Elnaghy, M. A., and Fadl-Allah, E. M. (1997). Isolation of *Pythium oligandrum* from Egyptian soil and its mycoparasitic effect on *Pythium ultimum* var. ultimum the damping-off organism of wheat. *Mycopathologia* 139:97. doi: 10.1023/A:1006836703594

Afgan, E., Baker, D., Batut, B., Van den beek, M., Bouvier, D., Čech, M., et al. (2018). The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544. doi: 10.1093/nar/gky379

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638

Aspeborg, H., Coutinho, P., Wang, Y., Brumer, H., and Henrissat, B. (2012). Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evol. Biol.* 12:186. doi: 10.1186/1471-2148-12-186

Benhamou, N., Bélanger, R. R., Rey, P., and Tirilly, Y. (2001). Oligandrin, the elicitin-like protein produced by the mycoparasite *Pythium oligandrum*, induces systemic resistance to Fusarium crown and root rot in tomato plants. *Plant Physiol. Biochem.* 39, 681–696. doi: 10.1016/s0981-9428(01)01283-9

Benhamou, N., Le Floch, G., Vallance, J., Gerbore, J., Grizard, D., and Rey, P. (2012). Pythium oligandrum: an example of opportunistic success. *Microbiology* 158, 2679–2694. doi: 10.1099/mic.0.061457-0

Benhamou, N., Rey, P., Picard, K., and Tirilly, Y. (1999). Ultrastructural and cytochemical aspects of the interaction between the mycoparasite *Pythium oligandrum* and soilborne plant pathogens. *Phytopathology* 89, 506–517. doi: 10.1094/PHYTO.1999.89.6.506

Berger, H., Yacoub, A., Gerbore, J., Grizard, D., Rey, P., Sessitsch, A., et al. (2016). Draft genome sequence of biocontrol agent *Pythium oligandrum* strain Po37, an oomycota. *Genome Announcements* 4, e00215–e00216. doi: 10.1128/genomeA.00215-16

Berry, L. A., Jones, E. E., and Deacon, J. W. (1993). Interaction of the mycoparasite *Pythium oligandrum* with other *Pythium species*. *Biocontrol Sci. Technol.* 3, 247–260.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Daly, P., Chen, Y., Zhang, Q., Zhu, H., Li, J., Zhang, J., et al. (2021). Pythium myriotylum is recovered most frequently from Pythium soft-rot infected ginger rhizomes in China. *Plant Disease*. Online ahead of print. doi: 10.1094/PDIS-05-21-0924-RE

Daly, P., Van Munster, J. M., Kokolski, M., Sang, F., Blythe, M. J., Malla, S., et al. (2017). Transcriptomic responses of mixed cultures of ascomycete fungi to lignocellulose using dual RNA-seq reveal inter-species antagonism and limited beneficial effects on CAZyme expression. *Fungal Genet. Biol.* 102, 4–21. doi: 10.1016/j.fgb.2016.04.005

De-Guzman, C. C., and Siemonsma, J. S. (1999). *Plant Resources of South East Asia: Spices*. Leiden: Backhuys Publishers.

Derevnina, L., Dagdas, Y. F., De La Concepcion, J. C., Bialas, A., Kellner, R., Petre, B., et al. (2016). Nine things to know about elicitins. *New Phytol.* 212, 888–895. doi: 10.1111/nph.14137

Doherty, A. J., and Jackson, S. P. (2001). DNA repair: how Ku makes ends meet. *Curr. Biol.* 11, R920–R924. doi: 10.1016/s0960-9822(01)00555-3

Doyle, J. J. (1990). Isolation of plant DNA from fresh tissue. *Focus* 12, 13–15.

FAOSTAT (2019). *Food and Agricultural Organisation of the United Nations Statistical Database*. Available online at: http://faostat.fao.org/ (accessed 24 July, 2021).

Faure, C., Veyssiere, M., Boelle, B., San Clemente, H., Bouchez, O., Lopez-Roques, C., et al. (2020). Long-read genome sequence of the sugar beet rhizosphere mycoparasite *Pythium oligandrum*. *G3 (Bethesda)* 10, 431–436. doi: 10.1534/g3.119.400746

Horner, N., Grenville-Briggs, L., and Van West, P. (2012). The oomycete *Pythium oligandrum* expresses putative effectors during mycoparasitism of *Phytophthora*

infestans and is amenable to transformation. *Fungal Biol.* 116, 24–41. doi: 10.1016/j.funbio.2011.09.004

Janků, M., Činčalová, L., Luhová, L., Jan, L., and Petřivalský, M. (2020). Biological effects of oomycetes elicitins. *Plant Protect. Sci.* 1, 1–8. doi: 10.17221/21/2019-pps

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., Mcanulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031

Kappel, L., Münsterkötter, M., Sipos, G., Escobar Rodriguez, C., and Gruber, S. (2020). Chitin and chitosan remodeling defines vegetative development and *Trichoderma biocontrol*. *PLoS Pathogens* 16:e1008320. doi: 10.1371/journal.ppat.1008320

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4

Köhl, J., Kolnaar, R., and Ravensberg, W. J. (2019). Mode of action of microbial biological control agents against plant diseases: relevance beyond efficacy. *Front. Plant Sci.* 10:845. doi: 10.3389/fpls.2019.00845

Kroon, L. P., Bakker, F. T., Van Den Bosch, G. B., Bonants, P. J., and Flier, W. G. (2004). Phylogenetic analysis of Phytophthora species based on mitochondrial and nuclear DNA sequences. *Fungal Genet. Biol.* 41, 766–782. doi: 10.1016/j.fgb.2004.03.007

Kushwaha, S. K., Vetukuri, R. R., and Grenville-Briggs, L. J. (2017). Draft genome sequence of the mycoparasitic oomycete *Pythium oligandrum* strain CBS 530.74. *Genome Announcements* 5, e346–e317. doi: 10.1128/genomeA.00346-17

Le, D. P., Smith, M., Hudler, G. W., and Aitken, E. (2014). Pythium soft rot of ginger: detection and identification of the causal pathogens, and their control. *Crop Protection* 65, 153–167. doi: 10.1016/j.cropro.2014.07.021

Le, D. P., Smith, M. K., and Aitken, E. A. B. (2016). An assessment of *Pythium spp*. associated with soft rot disease of ginger (*Zingiber officinale*) in Queensland, Australia. *Australasian Plant Pathol.* 45, 377–387. doi: 10.1007/s13313-016-0424-5

Le, D. P. (2016). *Characterization of Species Associated With Pythium Soft Rot of Ginger and Evaluation of Pythium oligandrum as a Biocontrol*. Ph.D thesis, Brisbane, QLD: University of Queensland

Lévesque, A. C., and De Cock, A. W. A. M. (2004). Molecular phylogeny and taxonomy of the genus Pythium. *Mycol. Res.* 108, 1363–1383. doi: 10.1017/s0953756204001431

Liang, D., Andersen, C. B., Vetukuri, R. R., Dou, D., and Grenville-Briggs, L. J. (2020). Horizontal gene transfer and tandem duplication shape the unique cazyme complement of the mycoparasitic oomycetes *Pythium oligandrum* and *Pythium periplocum*. *Front. Microbiol.* 11:581698. doi: 10.3389/fmicb.2020.581698

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Martin, F. N., and P. W. Tooley (2003). Phylogenetic relationships among *Phytophthora* species inferred from sequence analysis of mitochondrially encoded cytochrome oxidase I and II genes. *Mycologia* 95, 269–284. doi: 10.1080/15572536.2004.11833112

Mélida, H., Sandoval-Sierra, J. V., Diéguez-Uribeondo, J., and Bulone, V. (2013). Analyses of extracellular carbohydrates in oomycetes unveil the existence of three different cell wall types. *Eukaryot. Cell* 12, 194–203. doi: 10.1128/EC.00288-12

Mortazavi, A., Williams, B. A., Mccue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Meth* 5, 621–628. doi: 10.1038/nmeth.1226

Quinlan, R., Sweeney, M., Lo Leggio, L., Otten, H., Poulsen, J.-C., Johansen, K., et al. (2011). Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc. Natl. Acad. Sci. U S A.* 108, 15079–15084. doi: 10.1073/pnas.1105776108

Robideau, G. P., De Cock, A. W. A. M., Coffey, M. D., Voglmayr, H., Brouwer, H., Bala, K., et al. (2011). DNA barcoding of oomycetes with cytochrome c oxidase subunit I and internal transcribed spacer. *Mol. Ecol. Resources* 11, 1002–1011. doi: 10.1111/j.1755-0998.2011.03041.x

Singh, A. (2011). Management of rhizome rot caused by *Pythium*, *Fusarium* and *Ralstonia spp* in ginger (*Zingiber officinale*) under natural field conditions. *Ind. J. Agricultural Sci.* 81, 268–270.

Stirling, G. R., Turaganivalu, U., Stirling, A. M., Lomavatu, M. F., and Smith, M. K. (2009). Rhizome rot of ginger (*Zingiber officinale*) caused by *Pythium myriotylum* in Fiji and Australia. *Australasian Plant Pathol.* 38, 453–460. doi: 10.1071/ap09023

Thambugala, K. M., Daranagama, D. A., Phillips, A. J. L., Kannangara, S. D., and Promputtha, I. (2020). Fungi vs. fungi in biocontrol: an overview of fungal antagonists applied against fungal plant pathogens. *Front. Cell. Infect. Microbiol.* 10:604923. doi: 10.3389/fcimb.2020.604923

Tian, M., and Kamoun, S. (2005). A two disulfide bridge Kazal domain from Phytophthora exhibits stable inhibitory activity against serine proteases of the subtilisin family. *BMC Biochem.* 6:15. doi: 10.1186/1471-2091-6-15

Tomada, S., Sonego, P., Moretto, M., Engelen, K., Pertot, I., Perazzolli, M., et al. (2017). Dual RNA-Seq of *Lysobacter capsici* AZ78 - *Phytophthora infestans* interaction shows the implementation of attack strategies by the bacterium and unsuccessful oomycete defense responses. *Environ. Microbiol.* 19, 4113–4125. doi: 10.1111/1462-2920.13861

van der Plaats-Niterink, J. (1981). Monograph of the genus *Pythium*. *Stud. Mycol.* 21, 1–244.

Wei, L., Zhang, Q., and Zhou, D. (2019). Bio-Controlling Pythium oligandrum and its Application (translated). Patent Application Number CN201910757035.2. Beijing: China National Intellectual Property Administration.

Westermann, A. J., Barquist, L., and Vogel, J. (2017). Resolving host-pathogen interactions by dual RNA-seq. *PLoS Pathog* 13:e1006033. doi: 10.1371/journal.ppat.1006033

White, T. J., Bruns, T., Lee, S., and Taylor, J. (1990). "Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics," in *PCR Protocols*, eds M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (New York, NY: Academic Press), 315–322. doi: 10.1016/b978-0-12-372180-8.50042-1

Yang, K., Dong, X., Li, J., Wang, Y., Cheng, Y., Zhai, Y., et al. (2021). Type 2 Nep1-like proteins from the biocontrol oomycete *Pythium oligandrum* suppress *Phytophthora capsici* infection in solanaceous plants. *J. Fungi* 7:496. doi: 10.3390/jof7070496

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T. L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13:134. doi: 10.1186/1471-2105-13-134

You, X., Barraud, J., and Tojo, M. (2019). Suppressive effects of *Pythium oligandrum* on soybean damping off caused by *P. aphanidermatum* and *P. myriotylum*. *Ann. Report Kansai Plant Protection Soc.* 61, 9–13. doi: 10.4165/kapps.61.9

Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., et al. (2018). dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 46, W95–W101.

Conflict of Interest: LW, DZ, and QZ were co-authors on a patent application for *P. oligandrum* (Patent application number CN201910757035.2) relating to the biocontrol of plant diseases using the *P. oligandrum* GAQ1 isolate.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 22 December 2021
doi: 10.3389/fmicb.2021.772994

The Mevalonate Pathway Is Important for Growth, Spore Production, and the Virulence of *Phytophthora sojae*

Xinyu Yang[1,2], Xue Jiang[1], Weiqi Yan[1], Qifeng Huang[1], Huiying Sun[1], Xin Zhang[3], Zhichao Zhang[4], Wenwu Ye[4], Yuanhua Wu[1,2], Francine Govers[5] and Yue Liang[1,2]*

[1]College of Plant Protection, Shenyang Agricultural University, Shenyang, China, [2]Liaoning Key Laboratory of Plant Pathology, Shenyang Agricultural University, Shenyang, China, [3]Hunan Plant Protection Institute, Hunan Academy of Agricultural Sciences, Changsha, China, [4]College of Plant Protection, Nanjing Agricultural University, Nanjing, China, [5]Laboratory of Phytopathology, Wageningen University & Research, Wageningen, Netherlands

OPEN ACCESS

Edited by:
Xiao-Ren Chen, Yangzhou University, China

Reviewed by:
Yonglin Wang, Beijing Forestry University, China
Yufeng Francis Fang, GreenLight Biosciences, United States
Julio Vega-Arreguín, National Autonomous University of Mexico, Mexico

***Correspondence:**
Xinyu Yang
yangxy@syau.edu.cn

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *09 September 2021*
Accepted: *01 December 2021*
Published: *22 December 2021*

Citation:
Yang X, Jiang X, Yan W, Huang Q, Sun H, Zhang X, Zhang Z, Ye W, Wu Y, Govers F and Liang Y (2021) The Mevalonate Pathway Is Important for Growth, Spore Production, and the Virulence of Phytophthora sojae. Front. Microbiol. 12:772994. doi: 10.3389/fmicb.2021.772994

The mevalonate (MVA) pathway in eukaryotic organisms produces isoprenoids, sterols, ubiquinone, and dolichols. These molecules are vital for diverse cellular functions, ranging from signaling to membrane integrity, and from post-translational modification to energy homeostasis. However, information on the MVA pathway in *Phytophthora* species is limited. In this study, we identified the MVA pathway genes and reconstructed the complete pathway in *Phytophthora sojae in silico*. We characterized the function of the MVA pathway of *P. sojae* by treatment with enzyme inhibitor lovastatin, deletion of the geranylgeranyl diphosphate synthase gene (*PsBTS1*), and transcriptome profiling analysis. The MVA pathway is ubiquitously conserved in *Phytophthora* species. Under lovastatin treatment, mycelial growth, spore production, and virulence of *P. sojae* were inhibited but the zoospore encystment rate increased. Heterozygous mutants of *PsBTS1* showed slow growth, abnormal colony characteristics, and mycelial morphology. Mutants showed decreased numbers of sporangia and oospores as well as reduced virulence. RNA sequencing analysis identified the essential genes in sporangia formation were influenced by the enzyme inhibitor lovastatin. Our findings elucidate the role of the MVA pathway in *P. sojae* and provide new insights into the molecular mechanisms underlying the development, reproduction, and virulence of *P. sojae* and possibly other oomycetes. Our results also provide potential chemical targets for management of plant *Phytophthora* diseases.

Keywords: *Phytophthora sojae*, the mevalonate pathway, lovastatin inhibitor, geranylgeranyl diphosphate synthase, spore production, virulence

INTRODUCTION

Oomycetes are a diverse group of fungus-like eukaryotic microorganisms that are phylogenetically distinct from fungi and classified within the stramenopila kingdom. Various types of reproductive spores are produced in the oomycete life cycle, asexual sporangia, zoospores, and sexual oospores, and pathogenic oomycetes rely on both structures to achieve infection (Judelson and Blanco, 2005).

There are more than 100 species of oomycetes that cause destructive diseases in agriculture, and they are considered the most destructive species of plant pathogens (Kroon et al., 2012; Kamoun et al., 2015). For example, *Phytophthora infestans* is a notorious oomycete pathogen that causes late blight in potato and tomato worldwide (Fry, 2008). *Phytophthora sojae* is another notorious species that causes root and stem rot disease in soybean, resulting in approximately $2 billion losses worldwide annually (Tyler, 2007). In addition, the extraordinary genetic plasticity of *Phytophthora* enables the pathogens to adapt rapidly to, and overcome, chemical or host resistance, which makes controlling the diseases more difficult (Fawke et al., 2015).

Isoprenoids are the most diverse secondary metabolites in eukaryotes. They are produced *via* two independent pathways, the mevalonate (MVA) pathway which is present in most organisms, and the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway which is specific for plants and green algae (Vranova et al., 2013). In the initial steps of the MVA pathway, acetyl-CoA is converted into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), then IPP and DMAPP condense to produce geranyl diphosphate and farnesyl diphosphate (FPP; **Figure 1**; Nes, 2011, Fabris et al., 2014). FPP feeds into different sub-branches of the MVA pathway as a universal precursor (**Figure 1**; Hemmi et al., 2003, Fabris et al., 2014). Among the many enzymes participating in the MVA pathway, 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase, HMGR) is of particularly interest (**Figure 1**). This enzyme functions in the rate-limiting catalytic step and is the target of statins, the widely used cholesterol-lowering drugs (Goldstein and Brown, 1990; Lorenz and Parks, 1990). Binding of statins, such as lovastatin, to HMGR can alter the enzyme conformation, leading to decreased sterol biosynthesis (Istvan and Deisenhofer, 2001; Stancu and Sima, 2001). The ability of statins to inhibit sterol biosynthesis has been verified in human pathogenic fungi and provides antifungal properties (Westermeyer and Macreadie, 2007; Liu et al., 2009; Bellanger et al., 2016).

The sterol biosynthesis sub-branch of the MVA pathway has been identified in animals, fungi, and land plants (**Figure 1**; Desmond and Gribaldo, 2009; Ruiz-Sola et al., 2016). In contrast, oomycete species in the genera *Phytophthora* and *Pythium* are sterol auxotrophs that utilize exogenous sterols from the environment or from host plants to support their growth and development (Marshall et al., 2001; Yousef et al., 2009). Genome analyses revealed that most enzymes of the sterol biosynthesis sub-branch are absent in *Phytophthora* species, which makes them unable to synthesize sterol independently (Tyler et al., 2006; Dahlin et al., 2017). However, two enzymes (ERG3 and DHCR7) that function in the last steps of sterol biosynthesis were reported to present in *Phytophthora* and *Pythium* species (Wang et al., 2021a). These two enzymes were proposed to modify the exogenous sterols, especially DCHR7 was effective in converting ergosterol into brassicasterol in *Phytophthora capsici* (Wang et al., 2021b). Conversely, *Saprolegnia parasitica* and *Aphanomyces euteiches* are sterol autotrophs (i.e., synthesis cholesterol derivatives and fucosterol, respectively) and the sterol biosynthesis sub-branch has been identified in these two

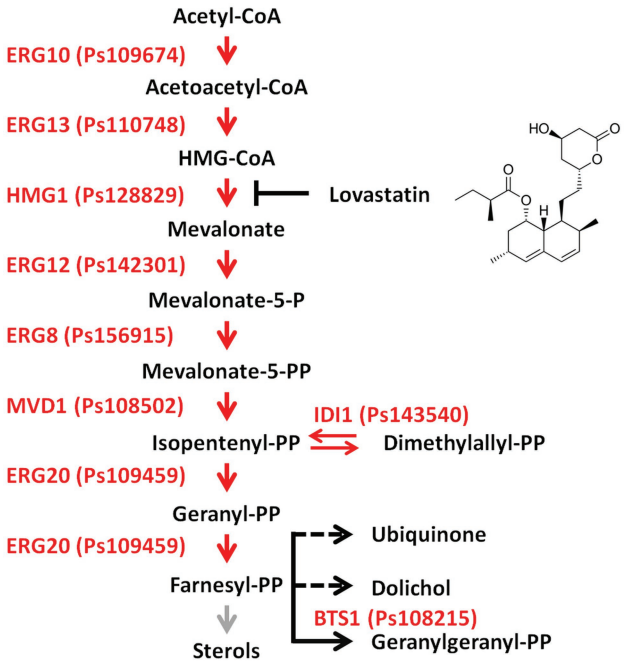


FIGURE 1 | Overview of the mevalonate (MVA) pathway in *Phytophthora sojae*. Enzymes and the corresponding gene ID of the MVA pathway in *P. sojae* are shown beside each arrow head. The main trunk of the MVA pathway is indicated by red arrows. The grey arrow indicated the deficient sterol biosynthesis sub-branch in *Phytophthora* species. The black arrow indicated the GGPP biosynthesis sub-branch, while the dashed arrow indicated that the sub-branches were not reported in *Phytophthora* species. The enzyme inhibitor lovastatin and chemical structure are also indicated on the right of the pathway. Acetyl-CoA, acetyl coenzyme A; Aceto-acetyl-CoA, aceto-acetyl coenzyme A; HMG-CoA, 3-hydroxy-3-methylglutaryl coenzyme A; Mevalonate-5-P, mevalonate phosphate; Mevalonate-5-PP, mevalonate diphosphate; Isopentenyl-PP, isopentenyl diphosphate; Dimethylallyl-PP, dimethylallyl diphosphate; Geranyl-PP, geranyl diphosphate; Farnesyl-PP, farnesyl diphosphate; and Geranylgeranyl-PP, geranylgeranyl diphosphate.

oomycete pathogens (Madoui et al., 2009; Warrilow et al., 2014). Moreover, genome and transcriptome analyses revealed that certain MVA pathway genes were potentially present in some *Phytophthora* species and the branch point enzyme geranylgeranyl diphosphate synthase (GGPS) in the terpenoid biosynthesis sub-branch of the MVA pathway was identified (Dahlin et al., 2017; Lerksuthirat et al., 2017).

GGPS utilizes FPP as precursor for the production of geranylgeranyl diphosphate (GGPP), the first product in terpenoid biosynthesis sub-branch (**Figure 1**; Hemmi et al., 2003; Beck et al., 2013). GGPS contains three highly conserved motifs: the G(Q/E) motif and two aspartate-rich motifs, designated the first aspartate-rich motif (FARM) and the second aspartate-rich motif (SARM; Hemmi et al., 2003). Based on the features of these conserved motifs, GGPSs are classified into three types: type I GGPS found in archaebacteria, type II GGPS found in bacteria and plants, and type III GGPS found in animals and fungi (Hemmi et al., 2003). In plants, GGPP serves as the precursor for photosynthetic pigments (carotenoids and chlorophylls) and phytohormones [gibberellins (GAs), abscisic acid, and

strigolactone] (Thabet et al., 2012; Ruiz-Sola et al., 2016). For example, the 10 functional GGPS genes identified in the *Arabidopsis* genome have differential spatiotemporal expression and subcellular localization (Okada et al., 2000; Beck et al., 2013). Among these GGPSs, mutation in GGPS1 affected chloroplast development while deletion of GGPS11 resulted in developmental defects in *Arabidopsis* (Ruppel et al., 2013; Ruiz-Sola et al., 2016). Overexpression of a sunflower GGPS gene in tobacco led to improved GA levels, and thereby an enhanced growth rate, early flowering, and increased seed yield (Tata et al., 2016).

In fungi, GGPS is not only essential for the production of isoprenoid-derived substances, such as antibiotics, the antitumor agent Taxol, and the antimalarial agent artemisinin, but also crucial for fungal development and virulence (Saikia and Scott, 2009; Singkaravanit et al., 2010a). For example, the *Saccharomyces cerevisiae* genome carries one GGPS gene, *BTS1*, that when deleted caused differential growth rates at lower temperatures (Jiang et al., 1995). The two different types of GGPS (*ggs1* and *ggs2*) isolated from filamentous fungi are responsible for production of GGPP or secondary metabolites (GA or paxilline), but the biological function has not been reported (Singkaravanit et al., 2010a). In the entomopathogenic fungus *Metarhizium anisopliae*, *ggs1* is constitutively expressed throughout fungal growth and is responsible for GGPP production (Singkaravanit et al., 2010b). Disruption of *ggs2* results in abolished helvolic acid production, delayed sporulation, and decreased toxicity to host insects (Singkaravanit et al., 2010a).

In *Phytophthora* spp., the MVA pathway has not been investigated and the function of the MVA pathway in development is not clear. In this study, we reconstructed the MVA pathway in *P. sojae* by identifying the whole set of genes and studied the effects of inhibiting the rate-limiting enzyme HMGR by lovastatin treatment. We also generated GGPS-deficient *P. sojae* mutants by CRISPR/Cas9-mediated deletion of the GGPS encoding gene *PsBST1*, analyzed their phenotypes, and compared the transcriptomes of wild-type (WT) *P. sojae* and a lovastatin-treated sample. This study shows that the MVA pathway plays essential roles in vegetative growth, development, reproduction, and virulence in *P. sojae* and uncovers a potential chemical target for sustainable management of soybean root rot caused by *P. sojae* and other *Phytophthora* diseases.

MATERIALS AND METHODS

Oomycete Strains and Culture Conditions

Wild-type oomycete strains *P. sojae* P6497, *P. infestans* T30-4, *P. capsici* LT263, and *Pythium ultimum* F18-6 were generously provided by Yuanchao Wang (Nanjing Agricultural University, Nanjing, China). All strains and *PsBTS1* deletion mutants were maintained on 10% V8 agar slants at 14°C in the dark at College of Plant Protection, Shenyang Agricultural University.

Gene Mining and Phylogenetic Analysis

The deduced amino acid sequences for genes related to the MVA pathway were downloaded from the yeast *S. cerevisiae* genome database[1] and used as BlastP queries to search for homologues from *P. sojae* (Joint Genome Institute, JGI). The homologues of *P. infestans*, *S. parasitica*, and *Py. ultimum* were searched from the Ensembl Genome database. The homologues of *P. capsici* were searched from NCBI GenBank. Each candidate amino acid sequence in *P. sojae* was blasted against the UniProt database to predict its potential functions. Based on a database derived from a global transcriptome investigation of *P. sojae* by 3′-tag gene expression analysis (Ye et al., 2011), the expression data of the nine genes in the MVA pathway were retrieved and transformed with log2 in this study. An expression heatmap was generated using Multiple Experiment Viewer with the hierarchical clustering method. GGPS proteins from different organisms were also downloaded from the National Center for Biotechnology Information (NCBI) database, and a phylogenetic tree was generated in MEGA 6.0 with the neighbor-joining method. The GGPS protein sequences were aligned using BioEdit software.

Inhibition Assay Under Statin Treatment

Lovastatin was purchased from Solarbio (Beijing, China) and a 20 mg/ml stock solution was prepared and stored at −20°C. The mycelial inhibition assays were preliminarily performed on V8 medium supplemented with different concentrations of lovastatin. Briefly, *P. sojae* was grown on V8 medium for 5 days; eight 5-mm mycelium plugs were transferred to a new plate; the plugs were flooded with 5 ml water containing various concentrations (0, 20, 40, and 80 μg/ml) of lovastatin; flooding was repeated every 40 min for 4 h; and then, plugs were incubated for another 5 h until sporangia formation. Zoospore suspensions (200 μl, 1×10^6 spores/ml) were gently mixed with lovastatin solution to final concentrations of 0, 20, 40, and 80 μg/ml, dropped into a glass slide, and incubated at 25°C in the dark for 40 min. The numbers of zoospores and encysted zoospores in a 10-μl suspension were counted; counts were repeated for a total of five times. After growth for 30 days on V8 medium supplied with the different concentrations of lovastatin, a 5-mm mycelium plug adjacent to the inoculation site was excised and the number of oospores was counted under a microscope (AXIO Scope A1; Zeiss, Oberkochen, Germany).

Plasmid Construction and CRISPR/Cas9-Mediated Gene Editing in *P. sojae*

We used two single guide RNAs (sgRNAs) and the HDR gene replacement method to delete *PsBTS1* (Fang and Tyler, 2016). We designed two efficient sgRNA sequences (sgRNA1-GCGTACATCAGCGAGCTGCC and sgRNA2-TCCAGAAGATGTCCAGCCCC) and inserted them to the pYF515 plasmid according to the method described by Fang et al. (2017). The 1,000-bp flanking sequences were ligated to the fluorescence gene eGFP (760 bp) on each side, which served as the donor DNA. The recombinant donor DNA sequences were inserted into pBlueScript SK II+ plasmid. All plasmids used in the CRISPR/Cas9 system were constructed as described

[1]https://www.yeastgenome.org/

by Fang et al. (2017). The full coding sequence of *PsBTS1* was amplified from the cDNA template and inserted into pTOReGFP for subcellular localization. All the plasmids were sequenced at the Beijing Genomics Institute (BGI, Shenzhen, China) before use. The CRISPR/Cas9-mediated gene replacement strategy and PEG-mediated protoplast transformation in *P. sojae* were used to obtain *PsBTS1* deletion mutants followed by the protocol presented by Fang and Tyler (2016). The gene deletion mutants were verified by PCR with genomic DNA to further confirm target gene deficiency; four individual deletion mutants from independent transformation were obtained and used as biological replicates for genotypic analysis subsequently. The PsBTS1 undeleted transformants were used as the control strains (CK) in the subsequent analyses.

Yeast Strains, Culture Conditions, and Transformation

The yeast gene deletion mutant strain BY4741 *bts1Δ* (*YPL069C*) and the WT strain BY4741 (*MATa his3Δ1*, *leu2Δ0 met15Δ0 ura3Δ0*) were purchased from Dharmacon (Cambridge, United Kingdom) and preserved in 30% glycerol. The full-length *PsBTS1* sequence was cloned into pYES2 (Invitrogen, Carlsbad, CA, United States) and transformed into the BY4741 *bts1Δ* mutant strain using a Yeastmaker Yeast Transformation System 2 kit (Takara, Dalian, China) following the manufacturer's protocol. The empty plasmid pYES2 was transformed into yeast WT strain BY4741 as well as the *bts1Δ* mutant, which served as the controls. Colonies grew on Synthetic Defined medium lacking uracil containing glucose were further screened by PCR. The transformed yeast cells were cultured overnight in YPD broth, then rinsed with water to remove YPD completely. The yeast cell solution was adjusted to an optical density at 600 nm of 0.2, then serially diluted 10 times. A 5-μl drop from each dilution was spotted into YPD medium without uracil, containing either glucose or galactose, and grown at 14, 25, or 30°C for 7 days.

Subcellular Localization by GFP Visualization

Phytophthora sojae transformants expressing PsBTS1-GFP fusion proteins and GFP were individually subcultured twice on V8 agar medium. Four mycelium plugs were transferred to a new plate, immersed in 10% V8 broth, and cultured for 2 days to obtain mycelia. The mycelia were rinsed twice with sterile distilled water to perform fluorescence microcopy with a confocal laser scanning microscope (FV3000; Olympus, Tokyo, Japan) using excitation/emission wavelengths of 488/515 nm.

Phenotype Characterization of *PsBTS1* Gene Deletion Mutants

All strains were subcultured twice on V8 agar before characterization analysis. The growth assay was carried out on Plich's medium using fresh cultures started on V8 agar and grown for 3–5 days at 25°C in the dark. Before the colony reached the edge of the plate, a mycelium plug 5 mm in diameter was taken and inoculated on Plich's medium. The plates were incubated at 25°C in the dark, then photographed after 10 days of incubation, and the mycelial growth and distance between branches were measured. The sporangia formation of *PsBTS1* mutants was induced by repeatedly flooding mycelial plugs with sterile water, the oospores numbers of *PsBTS1* mutants on a 5-mm mycelium plug were determined by visual counting under microscope.

Virulence Assays

The susceptible soybean cultivar Hefeng 47 was grown in vermiculite at 25°C for 4 days in the dark; the etiolated soybean seedlings were used to assess the virulence of *P. sojae* strains. Zoospore preparation was performed as described by Hua et al. (2008). Zoospore suspensions (5 μl, 20 zoospores/μl) of *PsBTS1* mutants were inoculated directly into hypocotyls of soybean seedlings, or zoospore suspensions were gently mixed with lovastatin solution to final lovastatin concentrations of 0, 20, 40, and 80 μg/ml and then inoculated on hypocotyls of soybean seedlings. The inoculated etiolated soybean seedlings were maintained at 25°C in the dark. The soybean seedlings were photographed and evaluated at 48 hpi. The epidermal cells from inoculation sites were excised at 12 and 24 hpi to explore the impeded virulence. All assays were performed three times.

Transcriptional Analysis of Lovastatin-Treated *P. sojae* Sample

To better understand the effects of MVA pathway on gene expression, RNA-seq was conducted using BGISEQ-500 sequencing platform at BGI. Based on our results, lovastatin (80 μg/ml)- and water-treated *P. sojae* strains were used for RNA-seq with three biological replicates per treatment. Trimmomatic was used to remove adaptors and low-quality reads from the raw data (Bolger et al., 2014); SOAP2 (Li et al., 2009) was used to map the filtered clean reads to the reference genome (*Phytophthora sojae* genome 1.0). The transcriptome data were assembled using StringTie (Pertea et al., 2016) to generate new transcripts, from which gene expression was calculated as the fragments per kilobase per million mapped reads. Differential expression analysis was completed using DESeq2 (Anders and Huber, 2010), where a fold change (treated expression/untreated expression) ≥ 2 and value of $p \leq 0.05$ were considered to indicate differential expression. Scatter plots were generated using the R base plot function. The enrichment analysis of GO terms and KEGG pathways from the JGI[2] was accomplished using the clusterProfiler package in R (Yu et al., 2012) with the hypergeometric distribution test and corrected value of p to enrich significant gene functions and pathways.

RNA Extraction and qRT-PCR Gene Expression Analysis

The *P. sojae* WT strain and a representative *PsBTS1* deletion mutant (T4) were grown on 10% V8 agar plates for 3 days at 25°C in the dark. Five mycelia plugs from the edges of colonies

[2]https://jgi.doe.gov/

were transferred to new plates and cultured in 10% V8 broth. The WT strain was treated with lovastatin (80 μg/ml). The mycelia were collected after 4 days of growth at 25°C in the dark for RNA extraction. RNA was extracted using a PureLink RNA Mini Kit (Thermo Fisher, Shanghai, China) and the quality was determined by 1.0% agarose gel electrophoresis. Total RNA (1–2 μg) was used to synthesize the first cDNA strand with oligo(dT) primers and an M-MLV Reverse Transcriptase Kit (Thermo Fisher, Shanghai, China). The primers for qRT-PCR were designed using Primer3 with an amplicon size range of 150–200 bp. SYBR green qRT-PCR assays were performed on an ABI StepOne system (Thermo Fisher, Foster City, United States) to evaluate expression levels of selected genes. Relative gene expression was calculated using the conserved *actinA* (accession number Ps108986) levels as internal controls (Hua et al., 2013; Yang et al., 2013). Means and SDs were calculated using data from three replicates. All assays were carried out following the manufacturers' protocols.

Statistical Analysis

All of the data were analyzed for statistical significance ($p < 0.05$) using GLM procedure for ANOVA or the Student's *t* test with SAS software (version 9.1; SAS Institute Inc., Cary, NC, United States).

RESULTS

The HMGR Inhibitor Lovastatin Reduces Growth, Zoosporulation, and Virulence of *P. sojae*

To elucidate the role of the MVA pathway, *P. sojae* was treated with lovastatin, a specific inhibitor of the rate-limiting enzyme HMGR (**Figure 1**). In *S. cerevisiae*, 10 μg/ml is sufficient to block the MVA pathway (Lorenz and Parks, 1990). In *P. sojae*, we observed that the growth rate decreased with increasing concentrations of lovastatin (**Figures 2A,D**) and the highest concentration that was tested, i.e., 80 μg/ml, mycelial growth was reduced with 31.9%. Similar inhibitory effects of lovastatin were observed in *P. infestans*, *P. capsici*, and *Py. ultimum* (**Supplementary Figure S1**). These results indicate that interruption of the MVA pathway by lovastatin treatment impairs the vegetative growth of *P. sojae* and closely related oomycetes.

To further assess the effects of the MVA pathway on the asexual and sexual development of *P. sojae*, sporangia and oospore formation were quantified in cultures grown in the presence of increasing concentrations of lovastatin (**Figures 2B,C**). Lovastatin treatment significantly inhibited sporangia and oospore formation (**Figures 2E,F**). Specifically, sporangia numbers were reduced by 67.7 and 93.3% at lovastatin concentrations of 20 and 40 μg/ml, respectively, while no sporangia were observed at 80 μg/ml (**Figures 2B,E**). Oospore numbers decreased by 38.4, 46.2, and 64.5%, respectively (**Figures 2C,F**). Lovastatin treatment also increased the zoospore encystment rate by 8.4-fold (**Figure 2G**).

In order to assess the effect of lovastatin on the virulence of *P. sojae*, zoospores were treated with different concentrations of lovastatin and subsequently used for inoculation of soybean hypocotyls. At 48 h post-inoculation (hpi), the length of lesions caused by zoospores treated with 80 μg/ml lovastatin decreased by 26.6% compared to mock-treated zoospores (**Figures 3A,B**). At 12 and 24 hpi, the zoospores without lovastatin treatment were able to germinate and penetrate epidermal cells, whereas the lovastatin-treated zoospores were unable to germinate (**Figure 3C**). The multiple effects of lovastatin, such as increased zoospore encystment rate, inhibited zoospores germination on host surface and mycelial extension, suggested that lovastatin reduces the virulence of *P. sojae* in this study.

Mining the *P. sojae* Genome for Genes Involved in the MVA Pathway

The deduced amino acid sequences of genes involved in the MVA pathway were downloaded from the *S. cerevisiae* genome database and used as BlastP queries to mine for potential homologues in the *P. sojae* genome. In total, nine homologues that could construct a complete MVA pathway were found (**Figure 1**; **Supplementary Table S1**). These genes showed pronounced similarity to homologues in *S. cerevisiae*, ranging from 29.5 to 51.4% at the amino acid level. In *S. cerevisiae*, most enzymes involved in the catalytic steps of the MVA pathway are encoded by a single gene, but the catalyzing enzyme HMGR is encoded by a pair of genes (*HMG1*/*HMG2*). In *P. sojae*, all homologues in the MVA pathway were identified, but only one HMGR gene was present (**Supplementary Table S1**). Furthermore, the *P. sojae* genes also showed high similarity to the identified homologues in the MVA pathways of *S. parasitica*, *P. infestans* (Dahlin et al., 2017), and *P. capsici* and *Py. ultimum*, with protein sequence identities in the range of 41.5–70.8, 73.3–93.9, 70.2–93.9, and 49.6–84.0%, respectively (**Supplementary Table S2**). The MVA pathway genes in *P. sojae* were further predicted in UniProt, the most similar homologues were identified in other *Phytophthora* spp. The similarity of MVA pathway genes of *P. sojae* with best-hit homologues ranges from 75.4 to 94.5% at the protein level.

Expression analysis of the genes in MVA pathway was conducted at the different development and infection stages. Then, a clustered heatmap was performed and gene expression was clustered into three differential patterns (**Supplementary Figure S2**). Specially, three genes (Ps143540, Ps142301, and Ps110748) showed relatively low expression levels while two (Ps109674 and Ps128829) were expressed in most stages of development and infection. The remaining four genes showed dynamically differential expression during the different stages, in which the gene encoding the branch point enzyme GGPS (accession number Ps108215; *PsBTS1*) exhibited a more dynamic expression pattern than the other three genes, including the relatively higher expressed in mycelial and infection stages and the lower in sporangia, zoospores, cysts, and germinated cysts. Based on the expression patterns, we hypothesized that GGPS (especially *PsBTS1*) is involved in growth and infection processes of *P. sojae*.

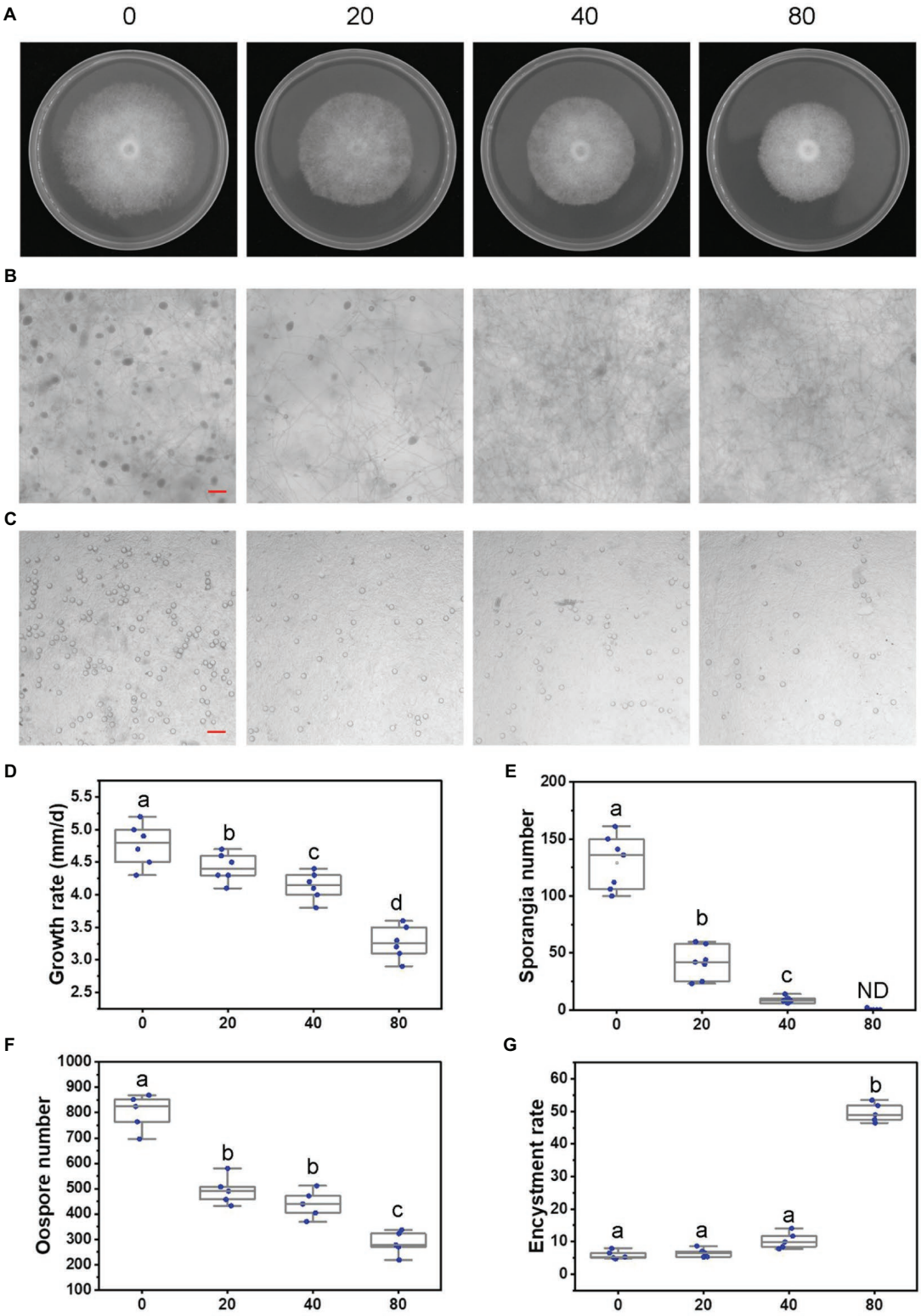


FIGURE 2 | Lovastatin suppresses growth, sporangia, and oospore formation and stimulates zoospore cysts in *P. sojae*. **(A)** *P. sojae* grown without and with lovastatin. **(B,C)** Microscopic visualization of sporangia in 15 days old cultures **(B)** and oospores in 30 days old cultures **(C)** grown without and with lovastatin. **(D)** *P. sojae* growth rates under treatment with various concentrations of lovastatin. **(E)** Sporangia counts on a 5-mm-diameter agar plug, ND, not detected. **(F)** Oospore counts on a 5-mm-diameter agar plug. **(G)** Encystment rates of zoospores harvested from 15 days old cultures grown without and with lovastatin. Lovastatin concentrations: 0, 20, 40, and 80 μg/ml; a, b, c, and d indicate significant differences ($p < 0.05$). Bar: 100 μm.

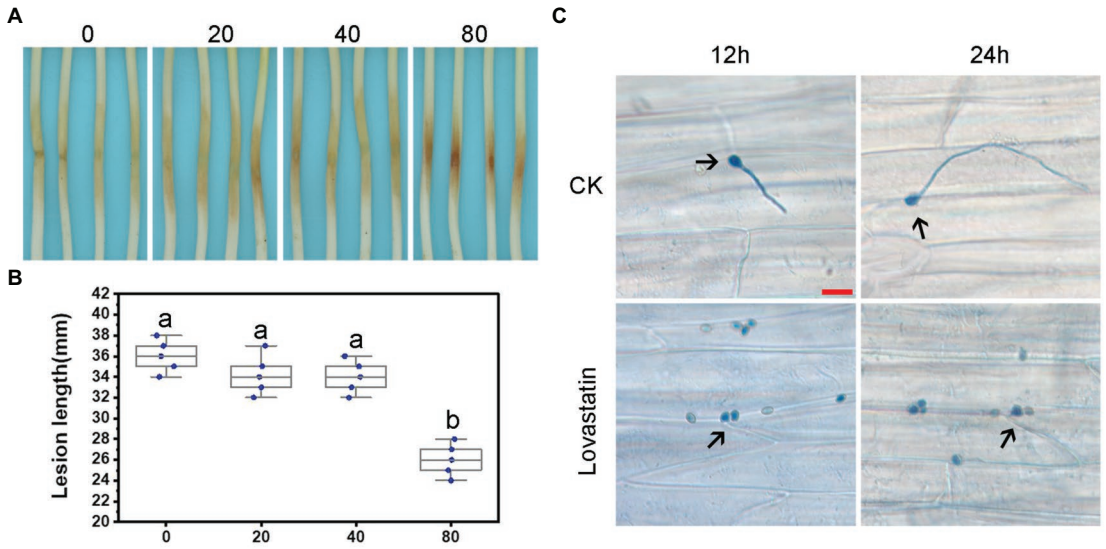


FIGURE 3 | Lovastatin treatment led to reduced virulence of *P. sojae*. Zoospore suspensions pretreated with various concentrations (0, 20, 40, and 80 μg/ml) of lovastatin were inoculated into hypocotyls of soybean cultivar Hefeng 47. **(A)** Photos and **(B)** lesion lengths were taken at 48 hpi. **(C)** The germinated cysts of untreated zoospores (CK) were observed on soybean epidermal cells at 12 and 24 hpi; lovastatin (80 μg/ml)-treated zoospores had not germinated at 12 or 24 hpi; and a and b indicate significant differences ($p < 0.05$). Bar: 25 μm.

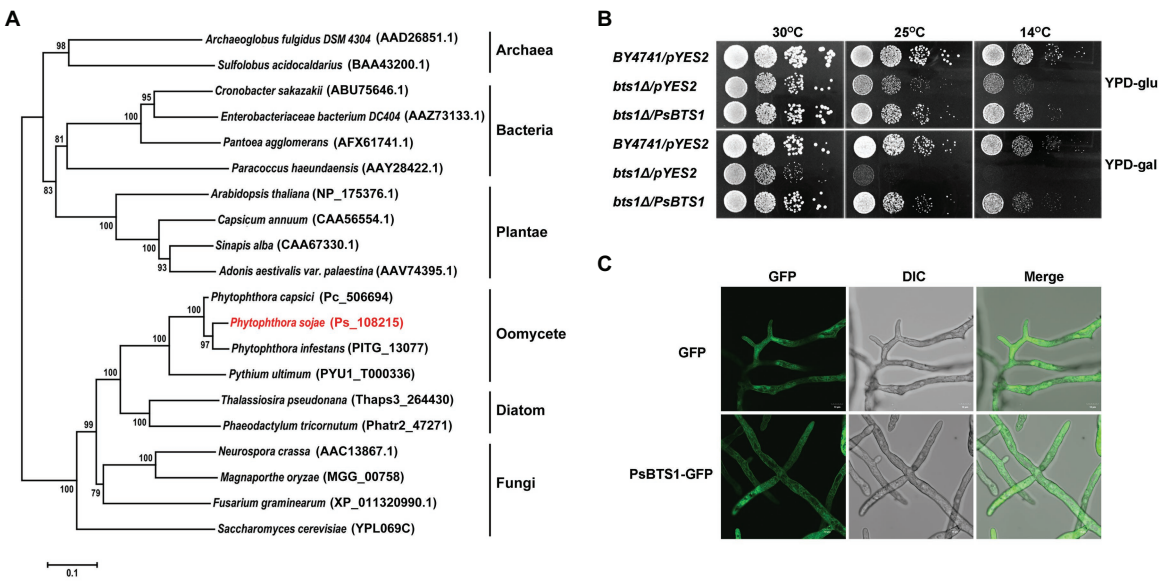


FIGURE 4 | Phylogenetic analysis, enzymatic activity, and subcellular localization of PsBTS1. **(A)** Phylogenetic analysis of GGPSs from different organisms. GGPS proteins from archaea, bacteria, plants, fungi, diatoms, and oomycetes were analyzed using MEGA6; their sequence IDs in the NCBI GenBank or genome databases are included. **(B)** PsBTS1 complements the *Saccharomyces cerevisiae bts1Δ* mutant growth defect at 14 and 25°C. Growth of *S. cerevisiae* strains *BY4741/pYES2*, *bts1Δ/pYES2*, and *bts1Δ/PsBTS1* was tested on YPD plates with glucose (top panel) and galactose (bottom panel) at 14, 25, and 30°C. **(C)** Subcellular localization of PsBTS1. GFP and PsBTS1-GFP fusion protein were expressed in *P. sojae* and mycelial were analyzed by fluorescence (GFP, 488/515 nm) and bright field (DIC) microscopy. Bar: 10 μm.

Characterization of the *P. sojae* GGPP Synthase *PsBTS1*

The gene encoding GGPP synthase is named *PsBTS1* after its homologue in *S. cerevisiae*. The genome sequence of *PsBTS1* comprises two introns, with a length of 108 and 107 bp, respectively. The 888 bp coding sequence (CDS) of *PsBTS1* encodes a protein of 295 amino acids in length. A phylogenetic tree based on PsBTS1 and its homologues from archaebacteria,

bacteria, fungi, plants, other oomycete species, and diatoms is comprised of three main clades, with archaebacteria clustering in one clade, bacteria and plants in another clade, and fungi, oomycetes, and diatoms together in the third clade (**Figure 4A**). The GGPS protein sequences of selected organisms were further used to identify the conserved motifs. The alignment revealed two aspartic-acid-rich motifs (FARM/SARM) and conserved G(Q/E) motifs (**Supplementary Figure S3**). Therefore, the GGPSs derived from oomycetes were categorized as type III, together with fungi and diatoms, especially as the FARM and SARM motifs differed from those of plants (**Supplementary Figure S3**).

Enzymatic Activity and Subcellular Localization of PsBTS1

To determine if *PsBTS1* encodes an active enzyme, we made use of *S. cerevisiae* cells in which the BTS1 gene (SGD:S000005990) is deleted. This temperature sensitive *bts1Δ* cells do not grow on galactose-containing medium at the non-permissive temperature. We transformed the expression plasmid pYES2 and the same plasmid carrying the CDS of *PsBTS1* into *bts1Δ* yeast cells and tested their growth on glucose-containing (suppression) or galactose-containing (induction) yeast extract–peptone–dextrose (YPD) medium at 14, 25, and 30°C (**Figure 4B**). All strains grew normally on

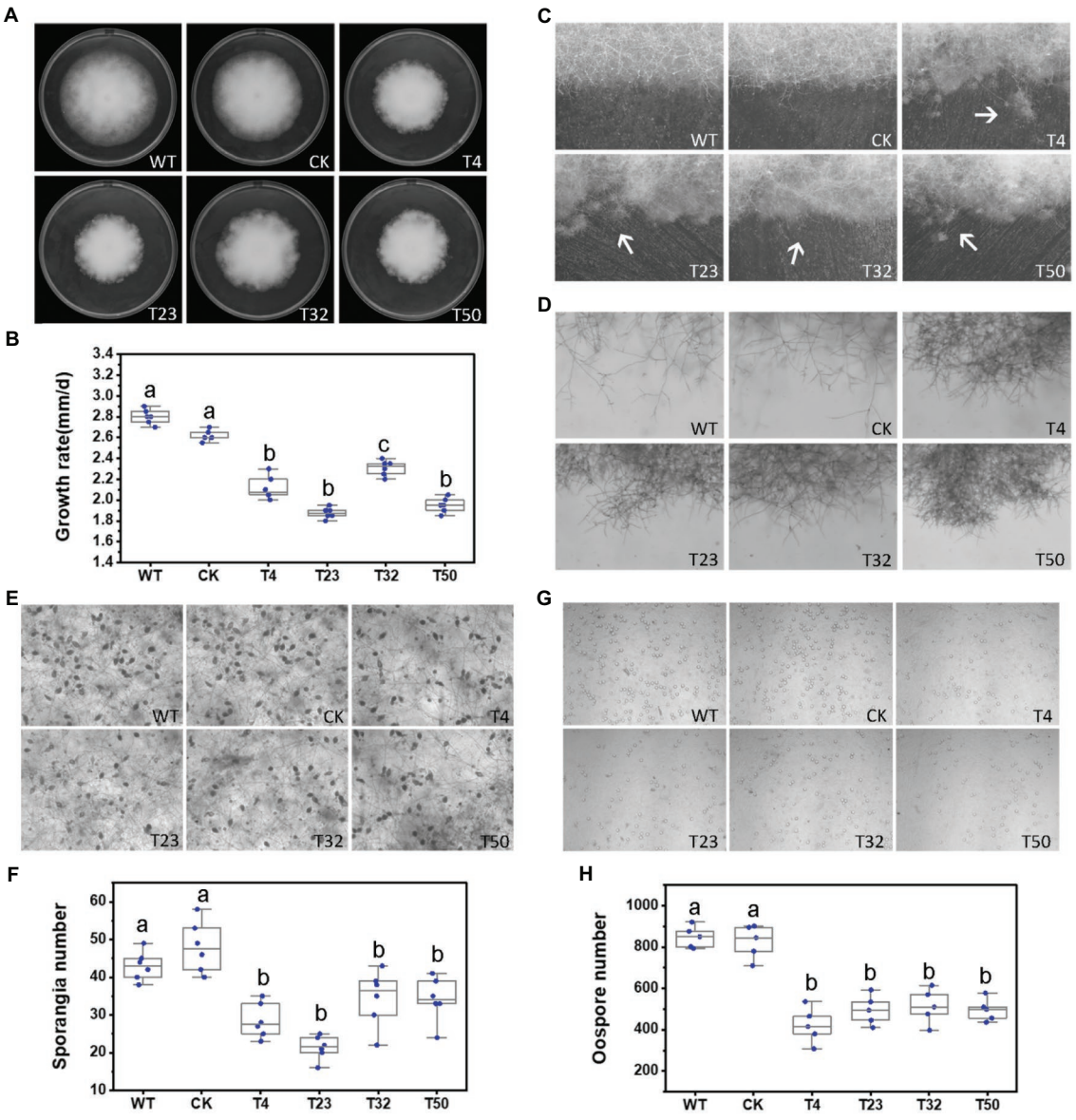


FIGURE 5 | Phenotypic analysis of *PsBTS1* deletion mutants. Comparison of wild-type (WT) *P. sojae* with a control transformant (CK) and four *PsBTS1* deletion mutants (T4, T23, T32, and T50). **(A)** Colony morphology and **(B)** growth rate on Plich medium. **(C)** Colony edges. **(D)** Mycelial morphology of satellite colonies. **(E)** Microscopic visualization and **(F)** quantification of sporangia. **(G)** Microscopic visualization and **(H)** quantification of oospores. In **(F)** and **(H)** a, b, and c indicate significant differences ($p < 0.05$). Bar: 100 μm.

both glucose- and galactose-containing YPD medium at the regular cultivation temperature of 30°C. However, the WT yeast strain BY4741 grew faster than the *bts1Δ*/pYES2 and *bts1Δ*/*PsBTS1* strains on glucose-containing medium at lower temperatures (14 and 25°C; **Figure 4B**). On galactose-containing YPD medium, the *bts1Δ*/*PsBTS1* strains formed colonies similar to the WT strains while the growth of *bts1Δ*/pYES2 strains was significantly inhibited (**Figure 4B**). Since these results demonstrate that *PsBTS1* can complement the growth defect of the yeast *bts1Δ* mutant, we conclude that PsBTS1 has the anticipated enzymatic activity and is thus a GGPP synthase.

To reveal the subcellular localization of PsBTS1, the CDS of *PsBTS1* was fused with green fluorescent protein (GFP) at the N-terminus, and the resultant fusion protein was expressed in *P. sojae*. A *P. sojae* strain with GFP expression was used as a control. The PsBTS1-GFP fusion protein and the control were driven by the constitutive ham34 promoter, which were highly expressed in *P. sojae* with the visual detection of strong green fluorescence. PsBTS1-GFP fusion protein and the control were found to be distributed in the cytoplasm (**Figure 4C**). In addition, the GFP signal observed in specific localization was not indicated in the magnification of either the PsBTS1-GFP mycelia and the control (**Supplementary Figure S4**), which suggested the expression of PsBTS1 was in the cytoplasm.

Phenotype Changes Caused by *PsBTS1* Deletion in the MVA Pathway

To investigate the role of *PsBTS1* in growth, reproduction, and virulence of *P. sojae*, we disrupted *PsBTS1* in *P. sojae* based on the CRISPR/Cas9-mediated homology-directed repair (HDR) method and replaced it with an exogenous GFP sequence. In this study, four deletion mutants (named T4, T23, T32, and T50) were obtained as the independent biological replicates and verified by PCR and quantitative reverse-transcription PCR (qRT-PCR) expression analyses (**Supplementary Figures S5B,C**). Editing of *PsBTS1* was observed based on the results of the junction PCR, while the spanning PCR and the internal PCR indicated an allele was replaced by exogenous GFP sequence (**Supplementary Figure S5B**). While the sequence of undeleted allele of PsBTS1 in the mutants was identical to that of WT (data not shown). All the mutants were heterozygous and showed decreased levels of *PsBTS1* transcript; no homozygous mutants were obtained also by making subcultures of single spores. The expression levels of two neighbor genes (Ps128815 and Ps128817) adjacent to *PsBTS1* were also estimated to avoid potential influence on the expression caused by target gene *PsPTS1* editing. Results indicated that the expression levels of the neighbor genes were similar among the *PsBTS1* deletion mutants and the WT strain (**Supplementary Figure S5D**). Morphological features of *PsBTS1* mutants were examined and compared to the WT and control (CK) strains at different developmental stages (**Figure 5**). Compared to WT and CK strains, the mutants exhibited abnormal and smaller colonies (**Figure 5A**). Specifically, the average rate of mycelial growth decreased by 24.1% (**Figure 5B**). Meanwhile, *PsBTS1* mutants exhibited rough and irregular colony edges where the aerial mycelial

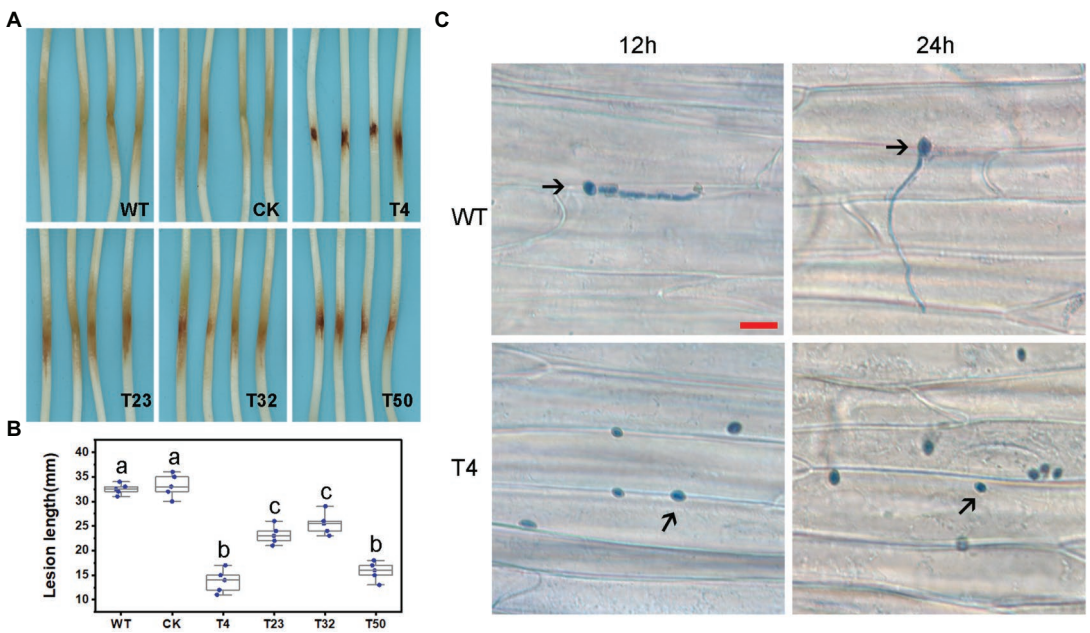


FIGURE 6 | PsBTS1 mutants showed reduced virulence. **(A)** Lesions and **(B)** lesion lengths on hypocotyls of etiolated soybean seedlings. Comparison of WT *P. sojae* with a CK and four PsBTS1 deletion mutants (T4, T23, T32, and T50), a, b, and c indicate significant differences ($p < 0.05$). **(C)** Microscopical examination of the epidermal cell surface at 12 and 24 hpi after inoculation with zoospores from WT *P. sojae* and *PsBTS1* KO T4. Bar: 25 μm.

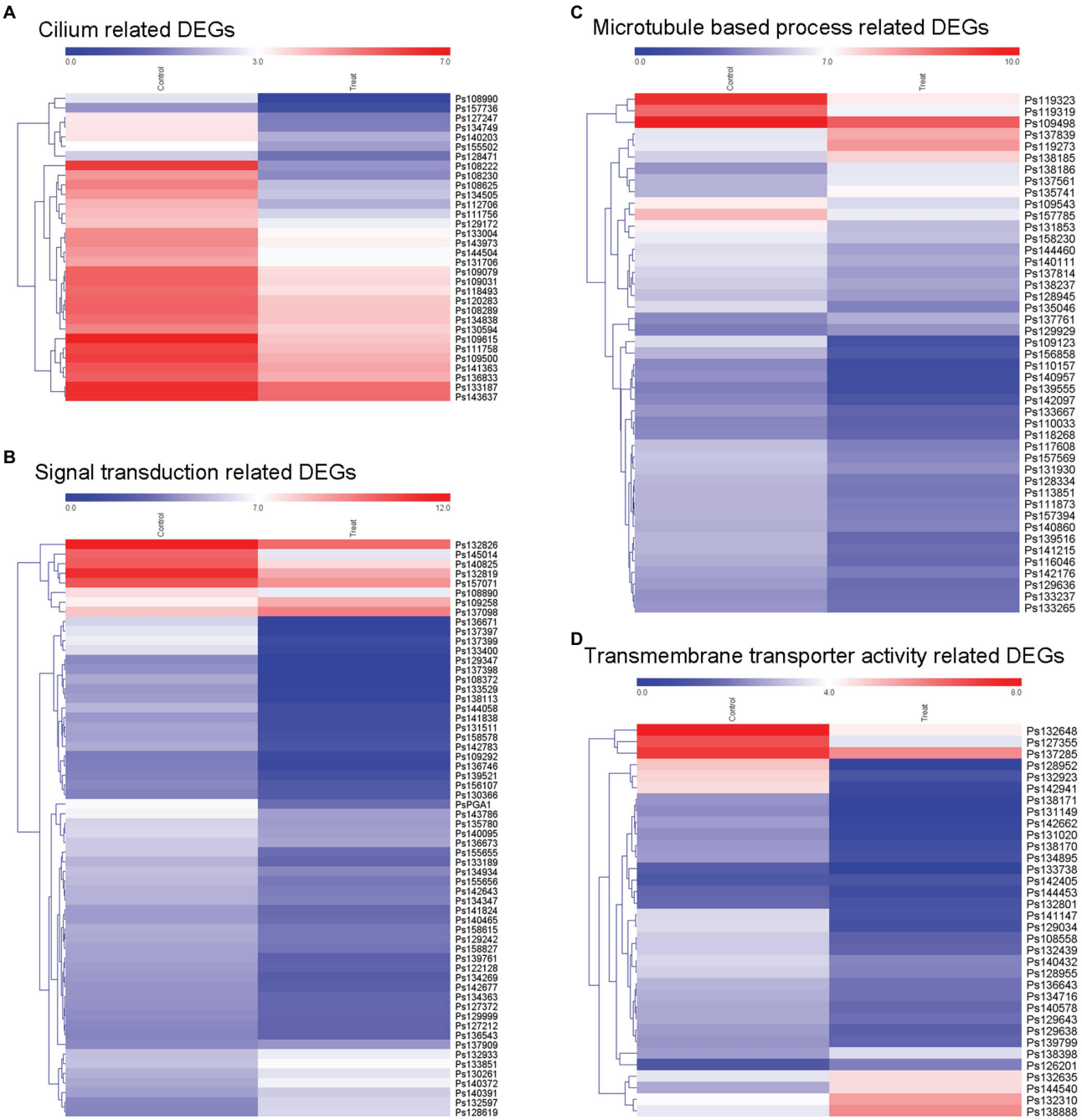


FIGURE 7 | Expression of different groups of genes associated with sporangia formation under lovastatin treatment. The GO terms of DEGs are associated with signal transduction, assembly of zoospore structures, and transport of necessary components. The expression changes of different groups of **(A)** cilium-related genes, **(B)** signal transduction-related genes, **(C)** microtubule-based process-related genes, and **(D)** transmembrane transporter activity-related genes were clustered and visualized in a heatmap.

tips attached to the surface of the growth medium and formed individual small satellite colonies (**Figure 5C**). The newly generated mycelia of *PsBTS1* mutants showed a more condensed hyphal density (**Figure 5D**), while the distance between branches was shorter in satellite colonies compared to WT and CK strains (**Supplementary Figure S6**). These results show that deletion of *PsBTS1* results in impaired colony morphology and vegetative growth of *P. sojae*.

We also assessed the functions of *PsBTS1* during asexual and sexual development by evaluating sporangia formation, zoospore encystment, and oospore production. The number of sporangia produced by *PsBTS1* mutants decreased compared

to the WT and CK (**Figures 5E,F**), while the number of oospores generated by *PsBTS1* mutants also declined (**Figures 5G,H**). Specifically, the sporangia numbers of the four *PsBTS1* mutants decreased 38.4, 55.8, 36.9, and 34.8%, respectively, compared to the WT and CK (**Figure 5F**). Meanwhile, the oospore numbers of the mutants decreased by 49.8, 37.3, 40.7, and 39.2% (**Figure 5H**). The encystment rate of *PsBTS1* mutants was not affected (data not shown). To further examine the infectivity of *PsBTS1* mutants, a virulence assay was performed on susceptible soybean seedlings. The lesions on soybean seedlings inoculated with *PsBTS1* mutant zoospores were reduced in length (**Figure 6A**); the lesion lengths caused by inoculation with the four mutants were decreased by 60.6, 26.5, 22.7, and 51.5% compared to the WT (**Figure 6B**). The WT zoospores germinated and penetrated into epidermal cells on the inoculation sites at 12 and 24 hpi, respectively, while the zoospores of the *PsBTS1* mutants did not germinate to cause infection (**Figure 6C**). Such results indicate that PsBTS1 plays important roles during asexual and sexual reproduction as well as in the virulence of *P. sojae*.

Gene Transcription Analysis

According to our results, the blocking of MVA pathway by lovastatin treatment or deletion of *PsBTS1* caused to reduced sporangia formation of *P. sojae*. Specifically, sporangia formation was significantly inhibited by lovastatin treatment (e.g., the concentration of 80 μg/ml) in this study. Such results suggest that sporangia formation is potentially related to the MVA pathway. To better elucidate the molecular mechanisms by which the MVA pathway mediates sporangia formation, transcriptome analysis was conducted by RNA-seq with *P. sojae* sporangia sample treated by lovastatin. In total, 4,018 differentially expressed genes (DEGs) were identified with a cut-off threshold of >2-fold change ($p < 0.05$), among which 1,393 genes were upregulated and 2,625 were downregulated (**Supplementary Figure S7A**). To predict the potential functions of the DEGs, Gene Ontology (GO) analysis was performed. The DEGs were primarily classified into biological process, cellular component, and molecular function (**Supplementary Figure S7B**). Specifically, the TOP 20 enriched GO terms (biological process) included cilium, signal transduction, microtubule-based process, and transmembrane transporter activity (**Supplementary Figure S7C**). The dynamic expression changes in key factors were revealed by the expression heatmap (**Figure 7**). The formation of sporangia involves complex regulation that requires signal transduction. In this study, all 37 identified cilium-related genes were downregulated (**Figure 7A**) and a total of 50 signaling pathway-related genes were identified with significant downregulation (**Figure 7B**). Microtubule-based process-related genes showed significant regulation, among which, 34 genes that were downregulated and 11 that were upregulated (**Figure 7C**). In total, 34 transmembrane transporter activity-related genes were detected, of which 28 were downregulated and six were upregulated (**Figure 7D**). In particular, dramatic expression changes in certain genes involved in signaling and cilia were identified, including the G protein alpha subunit (PsGPA1 and Ps108814) and phosphatase protein (PsCDC14 and Ps108222) in *P. sojae*, both of which were downregulated. Furthermore, the relative expression levels of *PsGPA1*, *PsCDC14*, and another six genes (*Ps158578*, *Ps142783*, *Ps109123*, *Ps139555*, *Ps109615*, and *Ps140432*) were validated by qRT-PCR in the lovastatin-treated sample and the representative KO mutant T4, all of which were downregulated, corroborating the RNA-seq results (**Supplementary Figures S8A,B**). The genes identified by RNA-seq analysis provide more information on the potential regulation of the MVA pathway in *P. sojae*.

DISCUSSION

Genome analysis indicated that many genes in sterol biosynthesis sub-branch of the MVA pathway are absent in *Phytophthora* spp., resulting in a sterol-auxotrophic life style (Tyler et al., 2006; Dahlin et al., 2017). Comparative analysis of sterol acquisition in sterol-auxotrophic and sterol-autotrophic oomycetes identified partial MVA pathway genes and the sterol biosynthesis sub-branch in *A. euteiches* and *S. parasitica* (Madoui et al., 2009; Dahlin et al., 2017). However, the complete MVA pathway in oomycetes has been little studied and its roles in development and virulence of these pathogens remain unknown. In this study, we blocked the MVA pathway by the enzyme inhibitor lovastatin which caused impaired growth, reproduction, and zoospore behavior of *P. sojae*. Subsequently, we identified a complete set of genes encoding enzymes that function in the MVA pathway in *P. sojae* and their homologues in related oomycete species. The enzyme inhibitor results and genes strongly suggest that the MVA pathway is ubiquitously conserved and functional in *Phytophthora* species. Deletion of the gene encoding the branch point enzyme PsBTS1 in *P. sojae* resulted in adverse effects on vegetative growth, reproductive development, and virulence similar to those caused by lovastatin treatment. Transcriptome analysis also further indicated that the MVA pathway potentially regulates growth, reproduction, and virulence characters in *P. sojae*.

The MVA pathway has been identified and is evolutionarily conserved in animals, plants, and fungi (Dimster-Denk et al., 1994; Fabris et al., 2014; Ruiz-Sola et al., 2016). The product of the initial steps in the MVA pathway, i.e., farnesyl diphosphate (FPP), feeds into different sub-branches to produce isoprenoids, sterols, ubiquinone, and dolichols, respectively (Fabris et al., 2014; Athanasakoglou et al., 2019). As an important metabolic pathway in cholesterol biosynthesis, the MVA pathway has been excavated as the pharmacological target of statins (Thompson and Taylor, 2017). The enzyme inhibitor statin targeted the rate-limiting enzyme HMGR to decrease cholesterol level in serum for hypercholesterolemia (Istvan and Deisenhofer, 2001; Stancu and Sima, 2001). The protein sequence, and the characteristics of the functional and binding domains in fungal HMGR share high similarity with the human homologue, which implies that statins also inhibit HMGR in fungal species (Bochar et al., 1999). Different formulations of statin were consistently found to inhibit growth of *Candida* species and *Aspergillus fumigatus* effectively

(Macreadie et al., 2006; Westermeyer and Macreadie, 2007). In addition to the observation that certain statins act as growth inhibitors, statins are also found to reduce biofilm production in the human pathogenic fungus *Candida albicans* (Liu et al., 2009), to increase the frequency of petite cells in *Candida glabrata* (Westermeyer and Macreadie, 2007), and to decrease germination in *Rhizopus oryzae* and cause melanin loss, reduced virulence, and increased susceptibility to oxidative stress in this species (Bellanger et al., 2016). Moreover, a decrease in ergosterol was confirmed in *Candida* species and *A. fumigatus* treated with statins, while treatments supplemented with ergosterol or cholesterol rescued statin-induced growth inhibition (Macreadie et al., 2006). These results indicated the similar effects of statins in sterol biosynthesis in human and fungi, the growth inhibition in fungi caused by statins is related to decreased sterol (Westermeyer and Macreadie, 2007; Bellanger et al., 2016).

Lovastatin treatment impeded development and virulence of *P. sojae*, which indicates a functional MVA pathway in *P. sojae*. *Phytophthora* zoospores are specific structure for short distance dispersal, which swim for hours with the help of flagella (Judelson and Blanco, 2005) and thereafter encyst on host epidermal cells to cause infection (Hua et al., 2008). Therefore, lovastatin was reported to impede zoospores dispersal due to the increase of encystment rate in this study. Moreover, the complete set of MVA pathway genes were found to be present in the genome of *P. sojae*, and all nine genes were found to be expressed. The different expression of these genes was supposed to play potential function in certain developmental or infection stages of *P. sojae*. Moreover, the homologues in the MVA pathway in *P. sojae* have also been identified in other *Phytophthora* species, which further indicates that the MVA pathway is conserved in *Phytophthora* species. However, the effects of statin as anti-*Phytophthora* agent were probably not identical to those in fungi because *Phytophthora* species specially lack sterol biosynthesis pathway (Tyler et al., 2006; Dahlin et al., 2017). The GGPS is a branch point enzyme in the terpenoid biosynthesis sub-branch of the MVA pathway (Alcaino et al., 2014). GGPSs are classified into three types based on the aspartate-rich motifs (Beck et al., 2013; Kato et al., 2016). The phylogenetic analysis revealed that oomycetes were closely related to two diatoms. This result is consistent with the fact that oomycetes and diatoms fall within the kingdom stramenopila and share a common ancestor (Tyler et al., 2006). Numerous GGPSs have been identified in plants and fungi that were found to localize in the cytoplasm or different organelles (Okada et al., 2000; Saikia and Scott, 2009; Beck et al., 2013). Thereby, the subcellular localization will suggest sub-functionalization by providing GGPP to specific tissues and even developmental stages (Beck et al., 2013; Coman et al., 2014). In plants, GGPS modulates development and growth by regulating biosynthesis of photosynthetic pigments and plant hormones (Ruppel et al., 2013; Ruiz-Sola et al., 2016; Tata et al., 2016; Zhou et al., 2017). For example, GGPSs in *A. thaliana* and *Oryza sativa* are involved in chlorophyll biosynthesis, and disruption leads to dwarf and leaf chlorosis phenotypes (Ruppel et al., 2013; Ruiz-Sola et al., 2016; Zhou et al., 2017).

Fungal GGPSs are involved in synthesis of diterpenes, such as crtE regulating astaxanthin production in *Xanthophyllomyces dendrorhous* (Alcaino et al., 2014), ggs2 affecting helvolic acid production in *M. anisopliae* (Singkaravanit et al., 2010a), and paxG associated with paxilline biosynthesis in *Penicillium paxilli* (Saikia and Scott, 2009). However, functional studies of GGPP in fungi are rare, with a few examples, such as deletion of the GGPS resulting in slow growth of *S. cerevisiae* under cold stress (Jiang et al., 1995) and deficiency of *ggs2* in *M. anisopliae* causing delayed sporulation and weak toxicity with no clear underlying mechanisms (Singkaravanit et al., 2010a). In our study, *PsBTS1* was identified and its encoded protein with enzymatic activity was localized in the cytoplasm. In addition, the monoallelic editing of *INF1* gene in *P. infestans* had been reported using CRISPR/Cas12a system recently, in which a copia-like element in the promoter region of *INF1* was inserted and presumably impeded the allele gene editing (Ah-Fong et al., 2021). According to current results in this study, *PsBTS1* editing resulted in defective mycelial growth, reduced sporangia and oospore numbers, and delayed germination of cysts on host surfaces, leading to decreased virulence in *P. sojae*. These phenomena were similar to the effects of lovastatin treatment. These results indicate the essential roles of the MVA pathway during the development, reproduction, and virulence of *P. sojae*.

Transitions between different stages in the life cycle contribute to the infection success of plant pathogenic oomycetes (Judelson and Blanco, 2005). Previous studies investigated transcriptional changes during the life cycles of *P. infestans* and *P. sojae* (Judelson et al., 2008, 2009; Ye et al., 2011). In particular, calcium-binding proteins, flagellar proteins, signaling proteins, and cation channel-encoding genes are upregulated at the sporangia stage (Judelson et al., 2008; Ah-Fong et al., 2017). In our study, sporangia formation was inhibited and sporangium-related gene expression was highly suppressed. For example, downregulated expression of genes encoding PsGPA1 and PsCDC14 was observed in this study. GPA1 regulates zoospore motility and virulence in *P. sojae* (Hua et al., 2008) and *P. infestans* (Latijnhouwers et al., 2004). CDC14 is expressed mainly in sporangia and regulates sporangia formation in *P. infestans* (Ah-Fong and Judelson, 2003, 2011). We hypothesize that the MVA pathway regulates growth, reproduction, and virulence by orchestrating dynamic changes at the transcriptome level in *P. sojae*.

In this study, we identified the complete MVA pathway in *P. sojae* and revealed its ubiquity in *Phytophthora* species. We demonstrated that the MVA pathway regulates growth, reproduction, and virulence in *P. sojae* through non-sterol pathway. Certain genes related to development and virulence were identified as being potentially affected by the MVA pathway in *P. sojae*. This research provides new clues about the molecular mechanisms involved in regulating development and virulence in *P. sojae*, which differ from those in fungi. Our findings may assist in the development of sustainable disease-management strategies and chemical targets for diseases caused by *P. sojae* and other *Phytophthora* species.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: PRJNA767514.

AUTHOR CONTRIBUTIONS

XY: conceptualization, formal analysis, investigation, writing – original draft, writing-review and editing, supervision, and funding acquisition. XJ, WYa, and QH: formal analysis, investigation, and visualization. HS, XZ, and ZZ: investigation. WYe and YW: formal analysis. FG: writing-review and editing. YL: conceptualization and writing-review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work is supported by the National Natural Science Foundation of China (grant nos: 31601586 and 31601615), the Scientific Research Foundation of Educational Department of Liaoning Province (LSNQN202015) and Starting Grants of Shenyang Agricultural University (880416060).

ACKNOWLEDGMENTS

We greatly thank Yuanchao Wang (College of Plant Protection, Nanjing Agricultural University) for providing *Phytophthora* strains and plasmid pYF515, Yuanhu Xuan and Xiaofeng Zhu (College of Plant Protection, Shenyang Agricultural University) for providing yeast expression vector pYES2 and Hefeng 47 soybean cultivar. We thank Suomeng Dong (College of Plant Protection, Nanjing Agricultural University) and Chenlei Hua (University of Tübingen) for helpful suggestion.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.772994/full#supplementary-material

REFERENCES

Ah-Fong, A. M. V., Boyd, A. M., Matson, M. E. H., and Judelson, H. S. (2021). A Cas12a-based gene editing system for *Phytophthora infestans* reveals monoallelic expression of an elicitor. *Mol. Plant Pathol.* 22, 737–752. doi: 10.1111/MPP.13051

Ah-Fong, A. M., and Judelson, H. S. (2003). Cell cycle regulator Cdc14 is expressed during sporulation but not hyphal growth in the fungus-like oomycete *Phytophthora infestans*. *Mol. Microbiol.* 50, 487–494. doi: 10.1046/j.1365-2958.2003.03735.x

Ah-Fong, A. M., and Judelson, H. S. (2011). New role for cdc14 phosphatase: localization to basal bodies in the oomycete *phytophthora* and its evolutionary coinheritance with eukaryotic flagella. *PLoS One* 6:e16725. doi: 10.1371/journal.pone.0016725

Ah-Fong, A. M., Kim, K. S., and Judelson, H. S. (2017). RNA-seq of life stages of the oomycete *Phytophthora infestans* reveals dynamic changes in metabolic, signal transduction, and pathogenesis genes and a major role for calcium signaling in development. *BMC Genomics* 18:198. doi: 10.1186/s12864-017-3585-x

Alcaino, J., Romero, I., Niklitschek, M., Sepulveda, D., Rojas, M. C., Baeza, M., et al. (2014). Functional characterization of the *Xanthophyllomyces dendrorhous* farnesyl pyrophosphate synthase and geranylgeranyl pyrophosphate synthase encoding genes that are involved in the synthesis of isoprenoid precursors. *PLoS One* 9:e96626. doi: 10.1371/journal.pone.0096626

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11:R106. doi: 10.1186/gb-2010-11-10-r106

Athanasakoglou, A., Grypioti, E., Michailidou, S., Ignea, C., Makris, A. M., Kalantidis, K., et al. (2019). Isoprenoid biosynthesis in the diatom *Haslea ostrearia*. *New Phytol.* 222, 230–243. doi: 10.1111/nph.15586

Beck, G., Coman, D., Herren, E., Ruiz-Sola, M. A., Rodriguez-Concepcion, M., Gruissem, W., et al. (2013). Characterization of the GGPP synthase gene family in *Arabidopsis thaliana*. *Plant Mol. Biol.* 82, 393–416. doi: 10.1007/s11103-013-0070-z

Bellanger, A. P., Tatara, A. M., Shirazi, F., Gebremariam, T., Albert, N. D., Lewis, R. E., et al. (2016). Statin concentrations below the minimum inhibitory concentration attenuate the virulence of *Rhizopus oryzae*. *J. Infect. Dis.* 214, 114–121. doi: 10.1093/infdis/jiw090

Bochar, D. A., Stauffacher, C. V., and Rodwell, V. W. (1999). Sequence comparisons reveal two classes of 3-hydroxy-3-methylglutaryl coenzyme A reductase. *Mol. Genet. Metab.* 66, 122–127. doi: 10.1006/mgme.1998.2786

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Coman, D., Altenhoff, A., Zoller, S., Gruissem, W., and Vranova, E. (2014). Distinct evolutionary strategies in the GGPPS family from plants. *Front. Plant Sci.* 5:230. doi: 10.3389/FPLS.2014.00230

Dahlin, P., Srivastava, V., Ekengren, S., McKee, L. S., and Bulone, V. (2017). Comparative analysis of sterol acquisition in the oomycetes *Saprolegnia parasitica* and *Phytophthora infestans*. *PLoS One* 12:e0170873. doi: 10.1371/journal.pone.0170873

Desmond, E., and Gribaldo, S. (2009). Phylogenomics of sterol synthesis: insights into the origin, evolution, and diversity of a key eukaryotic feature. *Genome Biol. Evol.* 1, 364–381. doi: 10.1093/gbe/evp036

Dimster-Denk, D., Thorsness, M. K., and Rine, J. (1994). Feedback regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 5, 655–665. doi: 10.1091/MBC.5.6.655

Fabris, M., Matthijs, M., Carbonelle, S., Moses, T., Pollier, J., Dasseville, R., et al. (2014). Tracking the sterol biosynthesis pathway of the diatom *Phaeodactylum tricornutum*. *New Phytol.* 204, 521–535. doi: 10.1111/NPH.12917

Fang, Y., Cui, L., Gu, B., Arredondo, F., and Tyler, B. M. (2017). Efficient genome editing in the oomycete *Phytophthora sojae* using CRISPR/Cas9. *Curr. Protoc. Microbiol.* 44:21A.1.1-21A.1.26. doi: 10.1002/cpmc.25

Fang, Y., and Tyler, B. M. (2016). Efficient disruption and replacement of an effector gene in the oomycete *Phytophthora sojae* using CRISPR/Cas9. *Mol. Plant Pathol.* 17, 127–139. doi: 10.1111/mpp.12318

Fawke, S., Doumane, M., and Schornack, S. (2015). Oomycete interactions with plants: infection strategies and resistance principles. *Microbiol. Mol. Biol. Rev.* 79, 263–280. doi: 10.1128/mmbr.00010-15

Fry, W. (2008). *Phytophthora infestans*: the plant (and R gene) destroyer. *Mol. Plant Pathol.* 9, 385–402. doi: 10.1111/j.1364-3703.2007.00465.x

Goldstein, J. L., and Brown, M. S. (1990). Regulation of the mevalonate pathway. *Nature* 343, 425–430. doi: 10.1038/343425a0

Hemmi, H., Noike, M., Nakayama, T., and Nishino, T. (2003). An alternative mechanism of product chain-length determination in type III geranylgeranyl diphosphate synthase. *Eur. J. Biochem.* 270, 2186–2194. doi: 10.1046/J.1432-1033.2003.03583.x

Hua, C. L., Meijer, H. J. G., de Keijzer, J., Zhao, W., Wang, Y. C., and Govers, F. (2013). GK4, a G-protein-coupled receptor with a phosphatidylinositol phosphate kinase domain in *Phytophthora infestans*, is involved in sporangia development and virulence. *Mol. Microbiol.* 88, 352–370. doi: 10.1111/MMI.12190

Hua, C., Wang, Y., Zheng, X., Dou, D., Zhang, Z., Govers, F., et al. (2008). A *Phytophthora sojae* G-protein α subunit is involved in chemotaxis to soybean isoflavones. *Eukaryot. Cell* 7, 2133–2140. doi: 10.1111/mpp.12279

Istvan, E. S., and Deisenhofer, J. (2001). Structural mechanism for statin inhibition of HMG-CoA reductase. *Science* 292, 1160–1164. doi: 10.1126/science.1059344

Jiang, Y., Proteau, P., Poulter, D., and Ferro-Novick, S. (1995). BTS1 encodes a geranylgeranyl diphosphate synthase in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 270, 21793–21799. doi: 10.1074/jbc.270.37.21793

Judelson, H., Ah-Fong, A., Aux, G., Avrova, A., Bruce, C., Cakir, C., et al. (2008). Gene expression profiling during asexual development of the late blight pathogen Phytophthora infestans reveals a highly dynamic transcriptome. *Mol. Plant Microbe Interact*. 21, 433–447. doi: 10.1094/MPMI-21-4-0433

Judelson, H. S., and Blanco, F. A. (2005). The spores of *Phytophthora*: weapons of the plant destroyer. *Nat. Rev. Microbiol.* 3, 47–58. doi: 10.1038/nrmicro1064

Judelson, H., Narayan, R., Ah-Fong, A., and Kim, K. (2009). Gene expression changes during asexual sporulation by the late blight agent *Phytophthora infestans* occur in discrete temporal stages. *Mol. Gen. Genomics.* 281, 193–206. doi: 10.1007/s00438-008-0407-5

Kamoun, S., Furzer, O., Jones, J. D., Judelson, H. S., Ali, G. S., Dalio, R. J., et al. (2015). The top 10 oomycete pathogens in molecular plant pathology. *Mol. Plant Pathol.* 16, 413–434. doi: 10.1111/mpp.12190

Kato, S., Takaichi, S., Ishikawa, T., Asahina, M., Takahashi, S., and Shinomura, T. (2016). Identification and functional analysis of the geranylgeranyl pyrophosphate synthase gene (crtE) and phytoene synthase gene (crtB) for carotenoid biosynthesis in *Euglena gracilis*. *BMC Plant Biol.* 16:4. doi: 10.1186/s12870-015-0698-8

Kroon, L. P., Brouwer, H., de Cock, A. W., and Govers, F. (2012). The genus *Phytophthora* anno 2012. *Phytopathology* 102, 348–364. doi: 10.1094/phyto-01-11-0025

Latijnhouwers, M., Ligterink, W., Vleeshouwers, V., Van West, P., and Govers, F. (2004). A G alpha subunit controls zoospore motility and virulence in the potato late blight pathogen *Phytophthora infestans*. *Mol. Microbiol.* 51, 925–936. doi: 10.1046/j.1365-2958.2003.03893.x

Lerksuthirat, T., Sangcakul, A., Lohnoo, T., Yingyong, W., Rujirawat, T., and Krajaejun, T. (2017). Evolution of the sterol biosynthetic pathway of *Pythium insidiosum* and related *Oomycetes* contributes to antifungal drug resistance. *Antimicrob. Agents Chemother.* 61, e02352–e02416. doi: 10.1128/AAC.02352-16

Li, R., Yu, C., Li, Y., Lam, T. W., Yiu, S. M., Kristiansen, K., et al. (2009). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25, 1966–1967. doi: 10.1093/bioinformatics/btp336

Liu, G., Vellucci, V. F., Kyc, S., and Hostetter, M. K. (2009). Simvastatin inhibits *Candida albicans* biofilm in vitro. *Pediatr. Res.* 66, 600–604. doi: 10.1203/PDR.0b013e3181bd5bf8

Lorenz, R. T., and Parks, L. W. (1990). Effects of lovastatin (mevinolin) on sterol levels and on activity of azoles in *Saccharomyces cerevisiae*. *Antimicrob. Agents Chemother.* 34, 1660–1665. doi: 10.1128/aac.34.9.1660

Macreadie, I. G., Johnson, G., Schlosser, T., and Macreadie, P. I. (2006). Growth inhibition of *Candida* species and *Aspergillus fumigatus* by statins. *FEMS Microbiol. Lett.* 262, 9–13. doi: 10.1111/j.1574-6968.2006.00370.x

Madoui, M. A., Bertrand-Michel, J., Gaulin, E., and Dumas, B. (2009). Sterol metabolism in the oomycete *Aphanomyces euteiches*, a legume root pathogen. *New Phytol.* 183, 291–300. doi: 10.1111/j.1469-8137.2009.02895.x

Marshall, J. A., Dennis, A. L., Kumazawa, T., Haynes, A. M., and Nes, W. D. (2001). Soybean sterol composition and utilization by *Phytophthora sojae*. *Phytochemistry* 58, 423–428. doi: 10.1016/s0031-9422(01).00219-9

Nes, W. D. (2011). Biosynthesis of cholesterol and other sterols. *Chem. Rev.* 111, 6423–6451. doi: 10.1021/cr200021m

Okada, K., Saito, T., Nakagawa, T., Kawamukai, M., and Kamiya, Y. (2000). Five geranylgeranyl diphosphate synthases expressed in different organs are localized into three subcellular compartments in *Arabidopsis*. *Plant Physiol.* 122, 1045–1056. doi: 10.1104/pp.122.4.1045

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11, 1650–1667. doi: 10.1038/nprot.2016.095

Ruiz-Sola, M. A., Coman, D., Beck, G., Barja, M. V., Colinas, M., Graf, A., et al. (2016). *Arabidopsis* geranylgeranyl diphosphate synthase 11 is a hub isozyme required for the production of most photosynthesis-related isoprenoids. *New Phytol.* 209, 252–264. doi: 10.1111/nph.13580

Ruppel, N. J., Kropp, K. N., Davis, P. A., Martin, A. E., Luesse, D. R., and Hangarter, R. P. (2013). Mutations in geranylgeranyl diphosphate synthase 1 affect chloroplast development in *Arabidopsis thaliana* (*Brassicaceae*). *Am. J. Bot.* 100, 2074–2084. doi: 10.3732/ajb.1300124

Saikia, S., and Scott, B. (2009). Functional analysis and subcellular localization of two geranylgeranyl diphosphate synthases from *Penicillium paxilli*. *Mol. Gen. Genomics.* 282, 257–271. doi: 10.1007/s00438-009-0463-5

Singkaravanit, S., Kinoshita, H., Ihara, F., and Nihira, T. (2010a). Cloning and functional analysis of the second geranylgeranyl diphosphate synthase gene influencing helvolic acid biosynthesis in *Metarhizium anisopliae*. *Appl. Microbiol. Biotechnol.* 87, 1077–1088. doi: 10.1007/s00253-010-2556-9

Singkaravanit, S., Kinoshita, H., Ihara, F., and Nihira, T. (2010b). Geranylgeranyl diphosphate synthase genes in entomopathogenic fungi. *Appl. Microbiol. Biotechnol.* 85, 1463–1472. doi: 10.1007/s00253-009-2171-9

Stancu, C., and Sima, A. (2001). Statins: mechanism of action and effects. *J. Cell. Mol. Med.* 5, 378–387. doi: 10.1111/j.1582-4934.2001.tb00172.x

Tata, S. K., Jung, J., Kim, Y. H., Choi, J. Y., Jung, J. Y., Lee, I. J., et al. (2016). Heterologous expression of chloroplast-localized geranylgeranyl pyrophosphate synthase confers fast plant growth, early flowering and increased seed yield. *Plant Biotechnol. J.* 14, 29–39. doi: 10.1111/pbi.12333

Thabet, I., Guirimand, G., Guihur, A., Lanoue, A., Courdavault, V., Papon, N., et al. (2012). Characterization and subcellular localization of geranylgeranyl diphosphate synthase from *Catharanthus roseus*. *Mol. Biol. Rep.* 39, 3235–3243. doi: 10.1007/s11033-011-1091-9

Thompson, P. D., and Taylor, B. (2017). Safety and efficacy of statins. *Lancet* 389, 1098–1099. doi: 10.1016/S0140-6736(17)30718-3

Tyler, B. M. (2007). *Phytophthora sojae*: root rot pathogen of soybean and model oomycete. *Mol. Plant Pathol.* 8, 1–8. doi: 10.1111/j.1364-3703.2006.00373.x

Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R., Aerts, A., et al. (2006). *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313, 1261–1266. doi: 10.1126/science.1128796

Vranova, E., Coman, D., and Gruissem, W. (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu. Rev. Plant Biol.* 64, 665–700. doi: 10.1146/annurev-arplant-050312-120116

Wang, W., Liu, X., and Govers, F. (2021a). The mysterious route of sterols in oomycetes. *PLoS Pathog.* 17:e1009591. doi: 10.1371/journal.ppat.1009591

Wang, W., Zhang, F., Zhang, S., Xue, Z., Xie, L., Govers, F., et al. (2021b). *Phytophthora capsici* sterol reductase PcDHCR7 has a role in mycelium development and pathogenicity. bioRxiv [Preprint]. doi: 10.1101/2021.04.17.440084

Warrilow, A. G. S., Hull, C. M., Rolley, N. J., Parker, J. E., Nes, W. D., Smith, S. N., et al. (2014). Clotrimazole as a potent agent for treating the oomycete fish pathogen *Saprolegnia parasitica* through inhibition of sterol 14 α-demethylase (CYP51). *Appl. Environ. Microbiol.* 80, 6154–6166. doi: 10.1128/aem.01195-14

Westermeyer, C., and Macreadie, I. G. (2007). Simvastatin reduces ergosterol levels, inhibits growth and causes loss of mtDNA in *Candida glabrata*. *FEMS Yeast Res.* 7, 436–441. doi: 10.1111/j.1567-1364.2006.00194.x

Yang, X., Zhao, W., Hua, C., Zheng, X., Jing, M., Li, D., et al. (2013). Chemotaxis and oospore formation in *Phytophthora sojae* are controlled by G-protein-coupled receptors with a phosphatidylinositol phosphate kinase domain. *Mol. Microbiol.* 88, 382–394. doi: 10.1111/MMI.12191

Ye, W., Wang, X., Tao, K., Lu, Y., Dai, T., Dong, S., et al. (2011). Digital gene expression profiling of the *Phytophthora sojae* transcriptome. *Mol. Plant-Microbe Interact.* 24, 1530–1539. doi: 10.1094/MPMI-05-11-0106

Yousef, L. F., Yousef, A. F., Mymryk, J. S., Dick, W. A., and Dick, R. P. (2009). Stigmasterol and cholesterol regulate the expression of elicitin genes in *Phytophthora sojae*. *J. Chem. Ecol.* 35, 824–832. doi: 10.1007/s10886-009-9653-1

Yu, G., Wang, L. G., Han, Y., and He, Q. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Zhou, F., Wang, C. Y., Gutensohn, M., Jiang, L., Zhang, P., Zhang, D., et al. (2017). A recruiting protein of geranylgeranyl diphosphate synthase controls metabolic flux toward chlorophyll biosynthesis in rice. *Proc. Natl. Acad. Sci. U. S. A.* 114, 201705689–201706871. doi: 10.1073/pnas.1705689114

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations,

or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 03 March 2022
doi: 10.3389/fmicb.2022.784686

The Transcription Factor *VpxlnR* Is Required for the Growth, Development, and Virulence of the Fungal Pathogen *Valsa pyri*

*Feng He[1,2], Alex-Machio Kange[3], Jie Yang[2], Jiaxin Xiao[2], Rongbo Wang[4], Lu Yang[2], Yifan Jia[1], Zheng Qing Fu[5], Yancun Zhao[1]*and Fengquan Liu[1]**

[1]Jiangsu Key Laboratory for Food Quality and Safety-State Key Laboratory Cultivation Base of Ministry of Science and Technology, Institute of Plant Protection, Jiangsu Academy of Agricultural Sciences, Nanjing, China, [2]College of Life Sciences, Anhui Normal University, Wuhu, China, [3]Department of Agriculture and Natural Resource, Bomet University College, Bomet, Kenya, [4]Fujian Key Laboratory for Monitoring and Integrated Management of Crop Pests, Fuzhou, China, [5]Department of Biological Sciences, University of South Carolina, Columbia, SC, United States

OPEN ACCESS

Edited by:
Vaibhav Srivastava, Royal Institute of Technology, Sweden

Reviewed by:
Hongyin Zhang, Jiangsu University, China
Wei Tang, Fujian Agriculture and Forestry University, China

***Correspondence:**
Yancun Zhao
zhaoyc27@126.com
Fengquan Liu
fqliu20011@sina.com

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *28 September 2021*
Accepted: *16 February 2022*
Published: *03 March 2022*

Citation:
He F, Kange A-M, Yang J, Xiao J, Wang R, Yang L, Jia Y, Fu ZQ, Zhao Y and Liu F (2022) The Transcription Factor VpxlnR Is Required for the Growth, Development, and Virulence of the Fungal Pathogen Valsa pyri. Front. Microbiol. 13:784686. doi: 10.3389/fmicb.2022.784686

Pears (*Pyrus* sp.) are widely cultivated in China, and their yield accounts for more than 60% of global pear production. The fungal pathogen *Valsa pyri* is a major causal agent of pear canker disease, which results in enormous losses of pear production in northern China. In this study, we characterized a Zn_2Cys_6 transcription factor that contains one GAL4 domain and a fungal-trans domain, which are present in VpxlnR. The *vpxlnR* gene expression was upregulated in the invasion stage of *V. pyri*. To investigate its functions, we constructed gene deletion mutants and complementary strains. We observed that the growth of the *vpxlnR* mutants was reduced on potato dextrose agar (PDA), Czapek plus glucose or sucrose compared with that of the wild-type strain. Additionally, *vpxlnR* mutants exhibited loss of function in fruiting body formation. Moreover, *vpxlnR* mutants were more susceptible to hydrogen peroxide (H_2O_2) and salicylic acid (SA) and were reduced in their virulence at the early infection stage. According to a previous study, VpxlnR-interacting motifs containing NRHKGNCCGM were searched in the *V. pyri* genome, and we obtained 354 target genes, of which 148 genes had Clusters of Orthologous Groups (COG) terms. PHI-BLAST was used to identify virulence-related genes, and we found 28 hits. Furthermore, eight genes from the 28 PHI-BLAST hits were further assessed by yeast one-hybrid (Y1H) assays, and five target genes, salicylate hydroxylase (VP1G_09520), serine/threonine-protein kinase (VP1G_03128), alpha-xylosidase (VP1G_06369), G-protein beta subunit (VP1G_02856), and acid phosphatase (VP1G_03782), could interact with VpxlnR *in vivo*. Their transcript levels were reduced in one or two *vpxlnR* mutants. Taken together, these findings imply that VpxlnR is a key regulator of growth, development, stress, and virulence through controlling genes involved in signaling pathways and extracellular enzyme activities in *V. pyri*. The motifs interacting with VpxlnR also provide new insights into the molecular mechanism of xlnR proteins.

Keywords: fungi, transcription factor, pear pathogens, virulence, stress

INTRODUCTION

Pear is the third most highly produced fruit in China. Valsa canker disease is one of the most destructive diseases in most orchards of northern China. The disease is caused by the fungal pathogen *Valsa pyri*, which belongs to Ascomycetes in the Valsaceae family (Sordariomycetes, Diaporthales; Yin et al., 2015). This fungus can infect pear trees from natural wound sites on the bark and then form cankers, which result in great yield loss or tree death [3, 4]. Although *Valsa mali* and *V. pyri* are similar species, they diverged 5 million years ago (Wang et al., 2014). *Valsa pyri* is a necrotrophic pathogen that can penetrate the phloem and xylem (Yin et al., 2015). Previously, it was shown that transcription factors (TFs), cell wall-degrading enzymes, and genes involved in nitrogen metabolism might be important for the virulence and growth of *V. pyri* strains (He et al., 2018; Xu et al., 2018). TFs, especially fungal-specific TFs, function as important regulators in fungal pathogens. However, very few studies have been conducted to investigate the roles of TFs in the pathogenesis of *V. pyri*.

Transcription factors can control the transcript levels of many target genes (Cho et al., 2013; He et al., 2016; Luo et al., 2016; Ishikawa et al., 2018; Oka et al., 2019). However, each gene can also be regulated by different TFs (Aro et al., 2001; Ishikawa et al., 2018). TFs and target genes constitute a network to regulate epigenetic modification, cell growth, cell differentiation, and stress responses (de Vries et al., 1999; Rauscher et al., 2006; Choi et al., 2009; Cho et al., 2013; Wu et al., 2018; Feng et al., 2020). To control gene expression, TFs generally possess one or more typical DNA-binding domains, and they are activated by themselves or other enzymes to bind promoter regions and induce mRNA transcription (Shelest, 2017; Ishikawa et al., 2018). Recently, as the full genomic sequences of more fungi became available, fungal TFs have been well characterized in different studies. According to a recent report, there are approximately 80 TF families in fungi, and many of them are fungus-specific TFs (Shelest, 2017). Fungus-specific TFs generally contain a typical fungal-trans domain, most of which contain other Zn_2Cys_6 clusters, and only a small portion of fungus-specific TFs contain C2H2 Zn fingers (MacPherson et al., 2006; Shelest, 2017). There are more than 100 fungus-specific TFs in most filamentous fungi (Shelest, 2017; He et al., 2018). Interestingly, many Zn_2Cys_6 TF orthologues exhibit various expression levels in different fungi, suggesting that they may have different roles in these fungi (He et al., 2018). Moreover, many TFs have been identified in many pathogenic fungi, and their roles and regulated genes have been well studied (Guo et al., 2011; Katz et al., 2013; Lu et al., 2014; He et al., 2016; Wu et al., 2018). For example, homeobox TFs are essential for conidiation and appressorium development (Kim et al., 2009). Hsf1 is a critical regulator of virulence traits (Veri et al., 2018), and VmSeb1 regulates development in *V. mali* (Wu et al., 2018), while VdMcm1 controls conidiation, microsclerotium formation, pathogenicity, and secondary metabolism (Xiong et al., 2016).

Although there are a large number of fungus-specific TFs, only a few are involved in virulence (Abe et al., 2007; Choi et al., 2009; Zhao et al., 2011; Cho et al., 2013; He et al., 2016; Wu et al., 2018). In previous studies, the fungus-specific TF AbPf2 and its orthologues were found to be involved in regulating fungal development, metabolism, and virulence in *Alternaria brassicicola*, *Verticillium dahlia*, and *Parastagonospora nodorum* (Cho et al., 2013; Luo et al., 2016; Rybak et al., 2017; Zhang et al., 2018). Due to the similar consensus sequences in the orthologues in various fungi, they generally have similar functions. Nevertheless, they also exhibit some unique roles in different isolates. For example, EBR1 orthologous gene FOXG_05408-knockout mutants in *Fusarium oxysporum* f. sp. lycopersici showed reduced virulence compared with the ebr1-deletion mutant in the PH-1 strain. These results indirectly proved the hypothesis that abundant Zn_2Cys_6 TFs may function in different processes and exhibit diverse functions in different fungal strains (Zhao et al., 2011; He et al., 2018). To control various aspects of fungal lifestyle, Zn_2Cys_6 TFs were previously reported to bind DNA motifs containing a CGG triplet (Cho et al., 2013; Luo et al., 2016; Raulo et al., 2016; Ishikawa et al., 2018). Thus, with high-throughput sequencing, such as transcriptome analysis or ChIP-seq, different motifs were found to interact with Zn_2Cys_6 TFs. In brief, Zn_2Cys_6 TFs bind a great number of genes to control fungal life.

XlnR, which contains one Zn_2Cys_6 cluster, has been characterized in several filamentous fungi and has vital roles in sugar metabolism in fungi (van Peij et al., 1998; MacPherson et al., 2006; Rauscher et al., 2006; Fujii et al., 2014). Its orthologues exhibit similar roles in several fungi. Thus, they can also regulate different genes in several fungi (Hasper et al., 2000; Calero-Nieto et al., 2007; Klaubauf et al., 2014; Llanos et al., 2019). The common function of XlnR in *Aspergillus* spp., *Trichoderma reesei*, *Fusarium species*, *Magnaporthe oryzae* (*Pyricularia oryzae*), and *Neurospora crassa* is to control xylanolytic and cellulolytic gene expression (Marui et al., 2002; Rauscher et al., 2006; Tamayo et al., 2008; Sun et al., 2012; Battaglia et al., 2013). XlnR regulates gene expression by binding to the CGGNTAAW motif as a monomer and by binding to the TTAGSCTAA motif as a dimer in *A. oryzae* (Ishikawa et al., 2018). These studies indicate that XlnR could function as a monomer and a dimer to control gene expression. Until now, the xlnR TF was only found to be involved in early infection by *F. graminearum*. However, its orthologous genes in *Valsa* species have not been identified. In this study, we discovered its roles in the pathogenicity of *V. pyri* and explored its molecular mechanism in controlling gene expression.

MATERIALS AND METHODS

VpxlnR Identification and Its Expression Pattern in the *Valsa pyri* Infection Stage

The TF open reading frame (ORF) sequence was obtained from the RNA-seq database (He et al., 2018), and the protein sequences were predicted using ORF finder.[1] The hypothetical protein was characterized using BlastP, and its orthologues,

[1]https://www.ncbi.nlm.nih.gov/orffinder/

including VmxlnR (KUI73112.1), hypothetical proteins (ROV95313.1, ROW10257.1, KAB5560083.1, XP_030979480.1, XP_016619423.1, and XP_001394612.2), TlxlnR (KAF3060643.1), BbxlnR (KAF1730510.1), MpxlnR (KLU82989.1), VlxlnR (KAG7119801.1), VdxlnR (XP_009650488.1), FoxlnR (RKL21210.1), and CgxlnR (KAF3801118.1), were acquired from the NCBI website. Similar to VpFSTF1, a phylogenetic tree was constructed using the hypothetical protein and its orthologous sequence by MEGA 7.0 (Kumar et al., 2016) and the hypothetical protein named VpxlnR. Furthermore, all protein sequences were submitted to the pfam database[2] by researching the protein domain, and later, the domains were drawn according to the research results. The transcript levels were evaluated as described previously (Kange et al., 2020).

Generation of Deletion Mutants for VpxlnR and Complementary Strains

DNA sequences approximately 2 kb upstream or downstream were acquired through BLASTn to *V. pyri* genomes, and primers 20 bp from a cassette containing the hygromycin phosphotransferase (*hph*) gene were designed. The primer pair of the cassette was also set with an 18–20 bp joint of upstream or downstream DNA sequences. We extracted Vp297 genomic DNA through the cetyltrimethylammonium bromide (CTAB) protocol (Umesha et al., 2016). Then, the upstream sequence, *hph* cassette and downstream sequence were amplified using the 1/2, 3/4, and 5/6 primer pairs (**Supplementary Table S5**), respectively. Then, the PCR products were purified using a PCR kit. Based on a previous study (He et al., 2016), the VpxlnR allele construct was amplified using the upstream sequence, *hph* cassette, and downstream sequence at a ratio of 1:3:1, and a 1/6 primer pair and primers were designed (**Supplementary Figure S2**; **Supplementary Table S5**). The VpxlnR allele was purified, and the purified product was transferred to the protoplast of the wild-type strain Vp297 using an improved polyethylene glycol (PEG)-mediated fungal transformation protocol (He et al., 2016). The VpxlnR deletion mutants were obtained on PDA medium after adding 50 mg/l hygromycin B, and approximately 120 primary transformants were acquired. To confirm whether the gene was deleted in these transformants, a partial DNA fragment from the ORF of VpxlnR was amplified using primer 7/8 primer pairs, and further, the *hph* fragment was amplified using 9/10 primer pairs (**Supplementary Figure S2**; **Supplementary Table S5**). Additionally, allele site replacement was ascertained using the 9/12 and 11/10 primer pairs (**Supplementary Figure S2**; **Supplementary Table S5**). In addition to genomic PCR, quantitative reverse transcription PCR (RT–qPCR) was carried out to evaluate the expression level of VpxlnR in the candidate mutants using the 15/16 primer pairs (**Supplementary Figure S2**; **Supplementary Table S5**).

We selected the vector pFL2 (Li et al., 2015a) as the overexpression plasmid, which was driven by the strong promoter RP27. The VpxlnR gene was amplified using primer pair 17/18, which contains a 20 bp sequence joint from the pFL2 plasmid, and then purified using a PCR kit (AxyPrep PCR Cleanup Kit, Suzhou, China). The resulting PCR product was constructed into XhoI-digested pFL2 through an one-step clone kit (Vazyme, Nanjing, China), and the recombinant vector was then transformed into *Escherichia coli* (DH5a cells). Later, the vector was extracted using a method described by a previous study (Kange et al., 2020). The plasmids were assessed by sequencing (GenScript, Nanjing, China), and the correct plasmids were used for further study. Then, the plasmid was transferred to protoplasts of the mutant m-56 by PEG-mediated transformation (He et al., 2016). The transformants were screened on PDA amended with 75 mg/L G418, and several positive transformants were further confirmed by genomic PCR and RT–qPCR.

Mycelial Growth and Fruiting Body Formation

Mycelial growth of the deletion mutants was characterized on PDA or Czapek media with different carbon sources, including glucose, sucrose, cellulose, sodium, pectin, and the control consisting of null sugar. Fruiting bodies were induced in the wild-type and complementary strains as reported in a previous study (He et al., 2016). The colony sizes of the wild-type Vp297, mutant and complementary strains grown on PDA media were measured at 36 h, and images were captured at the same time. Additionally, the colony sizes of the strains on Czapek media amended with glucose, sucrose, pectin, cellulose sodium, and null sugar were measured at 48 h, and later, these data were calculated using GraphPad prism 7.0. Every treatment was replicated on at least three agar plates.

Pathogenicity Assay

Pathogenicity assays were performed according to a previous study (Kange et al., 2019). Lesion development on inoculated leaves was observed daily, and images were captured at 3 and 5 days post-inoculation (dpi). Lesion lengths on inoculated branches were also observed daily, and images were captured at 2 and 4 dpi. Each experiment was duplicated with eight leaves and 13 branches, and the lesion size was calculated using GraphPad prism 7.0.

Host Mimic Stresses

Mycelial growth on PDA under H_2O_2 and SA stresses was used to assess the sensitivity of the mutants to host mimic stresses. The protocol used was similar to that used in a previous study (Kange et al., 2020). Colony sizes were measured at 36 h, and images were captured at the same time. Each experiment was repeated on at least three plates.

VpxlnR-Binding Promoter and Virulence-Related Gene Prediction

We first obtained all of the genes predicted in a previous study (Yin et al., 2015). Then, 2 kb predicted promoter sequences upstream of the initial gene codes were acquired. Later, according to the CGG triplet-containing DNA motif predicted by previous studies, we searched all predicted promoter sequences using

[2]http://pfam.xfam.org/

Regular Expression, and candidate promoters were further obtained (Cho et al., 2013). Moreover, candidate genes controlled by these promoters were acquired, and their DNA binding sites were also predicted using the Berkeley Drosophila Genome Project (BDGP, https://www.fruitfly.org/seq_tools/promoter.html). Additionally, their expression levels in the *V. pyri* infection stage were predicted according to *V. pyri* transcriptome analysis (He et al., 2018). The genes that were upregulated in the infection stage were selected, and later, their COG functional annotation was analyzed according to the gene annotation in *V. pyri* transcriptome analysis (He et al., 2018). Based on the gene annotation, we drew a COG term enrichment graph. Furthermore, to explore genes participating in virulence, all of the candidate genes were blasted to the PHI database, and genes with PHI hits were used for further identification.

Yeast One-Hybrid Assay

Based on COG annotation and PHI analysis, we selected eight genes, including salicylate hydroxylase (VP1G_09520), serine/threonine-protein kinase (VP1G_03128), alpha-xylosidase (VP1G_06369), G-protein beta subunit (VP1G_02856), acid phosphatase (VP1G_03782), serine/threonine-protein kinase KIN28 (VP1G_04075), serine/threonine-protein kinase GCN2 (VP1G_10966), and putative phosphotransferases (VP1G_03516), for the yeast one-hybrid (Y1H) assay. Of these genes, a 1,500–2,000 bp DNA sequence upstream of the start codon of each gene was amplified using genomic PCR, and then, the PCR products were purified using a PCR kit. The VpxlnR gene was ligated to a linearized pGADT7 vector. The promoter region of each gene was cloned into the pHIS2 vector, which was linearized by SmaI digestion. Positive clones were further confirmed by sequencing, and plasmids were obtained from the GenScript Company (Nanjing, China). The pAD::VpxlnR vector and the pHIS2::promoter vector were cotransformed into AH109 Gold yeast competent cells. Positive clone screening and confirmation were performed in accordance with a previous study (Kange et al., 2019).

Expression Levels of the Target Genes Controlled by VpxlnR

According to the Y1H results, we analyzed the transcript levels of eight genes (VP1G_06369, VP1G_03128, VP1G_03782, VP1G_09520, and VP1G_02856) that were directly controlled by VpxlnR and three genes (VP1G_04075, VP1G_03516, and VP1G_10966) with negative results of Y1H. The RT–qPCR protocol and analysis of the results were the same as those in a previous study (Kange et al., 2019). Each experiment was repeated at least three times.

RESULTS

Phylogenetic Analysis and Expression Pattern of VpxlnR in *Valsa pyri*

In previous studies, we found that many genes in *V. pyri* were annotated as xlnR homologues through RNA-Seq analysis (He et al., 2018), and its unigene (c14467) was reidentified through *de novo* assembly using transcriptome data. The gene ORF with a length of 2,625 bp was predicted. Alignment using the unigene and genomic gene sequences was performed. As a result, we found that the gene contains four introns; however, the genome assembly was missing a 12-bp sequence, which might cause the prediction using genomics and transcriptomics to vary (**Supplementary Figure S1**). The xlnR-like protein predicted using transcriptomics contains a Zn_2Cys_6 zinc cluster and fungal-trans domain and is highly similar to other xlnRs in other fungal species, especially *Valsa mali* (**Figure 1A**), so we named it VpxlnR. Moreover, the domains of VpxlnR orthologues were analyzed. In most selected fungi, orthologues have similar domains consisting of a Zn_2Cys_6 cluster and a fungal-trans domain, but several VpxlnR orthologues lost the Zn_2Cys_6 cluster in *Valsa malicola*, *Valsa sordida*, *V. dahlia*, *Colletotrichum gloeosporioides*, and *Aspergillus niger* (**Figure 1A**), which may be caused by alternative splices or natural selection. These results suggested that VpxlnR is a fungus-specific transcription factor.

To further study the transcript levels of *VpxlnR*, total RNA samples of mycelia and mycelial infection pear bark were acquired [0–48 h post-inoculation (hpi)], and RT–qPCR was performed. At the early infection stage, *VpxlnR* expression was upregulated until 6 h after inoculation, but its transcript levels were downregulated at 12 h after inoculation. Furthermore, its transcript level restored the upregulation at the late infection stage (**Figure 1B**). These results suggest that the expression of the *VpxlnR* gene can be induced by pear plants.

Generation of *VpxlnR* Deletion Mutants and Complementary Strains

To determine the functions of VpxlnR in *V. pyri*, two *VpxlnR* deletion mutants were generated using PEG-mediated transformation (**Supplementary Figure S2**). The deletion mutants were selected on PDA medium containing 50 mg/L hygromycin B. The *VpxlnR* gene in the two mutants was successfully replaced by a hygromycin-resistant cassette, which we named m-7 and m-56 (**Supplementary Figure S2A**). Furthermore, we constructed two complementary strains using protoplasts of the m-56 mutant as the target cells by PEG-mediated transformation, and the PCR amplification results showed that the *VpxlnR* gene was successfully amplified in these complementary strains (**Supplementary Figure S2A**). To examine the expression levels of *vpxlnR* in these mutants and complementary strains, RT–qPCR assays were performed, and the results showed that the expression of the *vpxlnR* gene was not detected in the two mutants, while the transcript levels of *vpxlnR* were normal in the two complementary strains (**Supplementary Figure S2B**). Taken together, these data indicate that we successfully acquired *vpxlnR*-knockout mutants and complementary stains of *V. pyri*.

VpxlnR Controls Mycelial Growth and Fruiting Body Formation

To characterize the morphological features of the *vpxlnR* mutants, we recorded mycelial growth on PDA, pear branch agar (PBA),

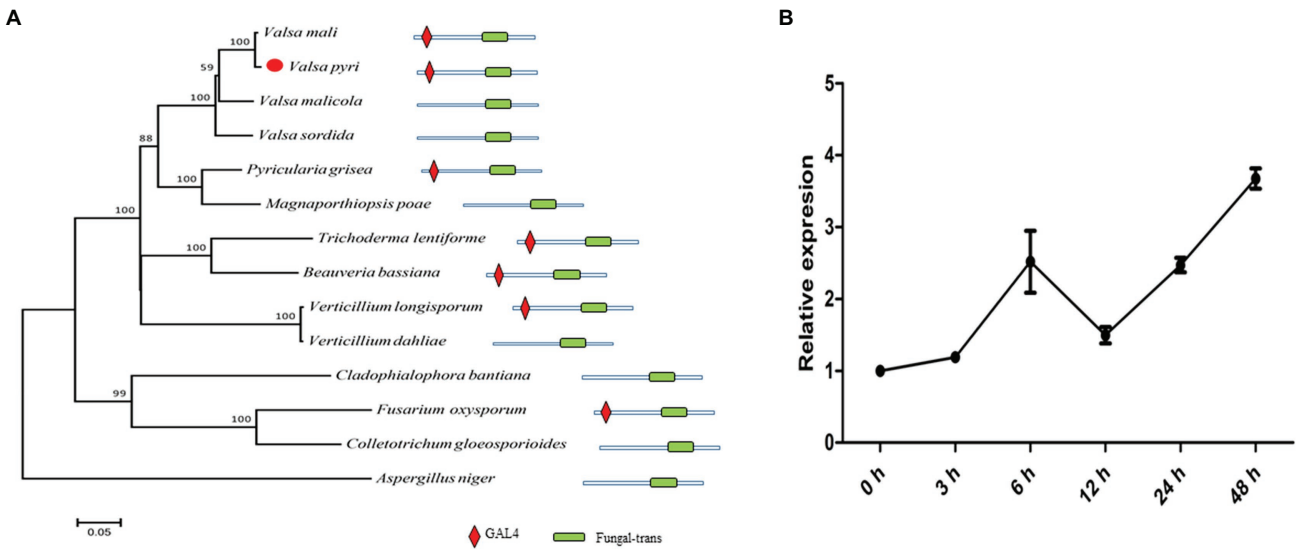


FIGURE 1 | VpxlnR characterization and its expression pattern during infection. **(A)** Phylogenetic tree of xlnR orthologues in different fungi. The VpxlnR gene sequence was obtained from *Valsa pyri* transcriptomics data, and sequences of xlnR orthologues were downloaded from the NCBI database. Sequence alignment was carried out using ClustalW, and a phylogenetic tree was generated *via* the neighbor-joining method using MEGA 7.0 software. The confidence values above the nodes were acquired from 1,000 bootstrap analyses. The domain architecture of VpxlnR and its orthologues were searched with the Pfam database. The domains were manually drawn according to the search results. **(B)** Transcript levels of *VpxlnR* during *V. pyri* infection. The relative expression levels of *Vpxl*nR in *V. pyri* were quantified by quantitative reverse transcription PCR (RT–qPCR) and normalized to the transcript levels of the *V. pyri* actin gene. Fold changes were calculated by the $2^{-\Delta\Delta Ct}$ approach. Each experiment was performed three times.

CM media, or Czapek media containing different carbon sources for fruiting body production. In the mutants cultured on both PDA, PBA, and CM media, 36 h later, the colony sizes of the *vpxlnR* mutants were significantly smaller than those of the wild-type (Vp297) and the *VpxlnR*-overexpressing complementary strains (**Figure 2A**). Because we added equal amounts of glucose to the three media, the growth ratios of the stains were similar on the three media. We speculate that the deletion of *VpxlnR* might weaken the ability of *V. pyri* to use carbon resources. To evaluate the abilities of *vpxlnR* mutants in the utilization of carbon resources, the mutants were grown on Czapek media with pectin, sucrose, glucose, and cellulose and with no sugar as the control. We found that the colony sizes of the mutants were obviously smaller than those of the wild-type or complementary strains (**Figure 2B**) on the five media. These results suggest that VpxlnR may determine glucose and sucrose assimilation, which results in a low growth ratio of the mutants. Furthermore, the mutants and other strains were cultured under light conditions. After 20 days, the mutants failed to produce fruiting bodies, while the wild-type and complementary strains produced fruiting bodies, indicating that VpxlnR affected fruiting body formation (**Figure 2C**). Taken together, these results demonstrated that VpxlnR plays important roles in the growth and development of *V. pyri*.

VpxlnR Regulates Virulence of *Valsa pyri*

To assess the effect of VpxlnR on virulence, mycelial agar plugs of the wild-type, mutant, and complementary strains were inoculated on detached pear leaves and branches. The diameters of the lesions on leaves were measured at 2, 3, 4, and 5 dpi. We found that the lesion sizes produced by the *vpxlnR* mutants were significantly smaller than those produced by the wild-type and complementary strains at 2–5 dpi (**Figures 3A,B**). Accordingly, the lesion length caused by the mutants on the inoculated branches was significantly smaller at each time point (2–6 dpi; **Figures 3C,D**). These results suggest that VpxlnR may play an important role in the determination of infection progression at the invasion stage on pear trees.

The *vpxlnR* Mutants Are Sensitive to Host Mimic Stress in a Concentration-Dependent Manner

Pathogens must overcome stresses, especially reactive oxygen species (ROS) or immune signalling chemicals such as salicylic acid (SA), from plants to invade successfully. Because the pathogenicity of the *vpxlnR* mutants was significantly reduced on pear branches and leaves, they may be more sensitive to host immune responses such as ROS burst and SA accumulation. To evaluate whether the mutants could resist host-derived stresses, we recorded the growth status on PDA media amended with different concentrations of H_2O_2 or SA. With increasing SA concentration, the colony size of all of the strains was reduced at 36 h post-inoculation (hpi) on PDA. Compared with the wild-type strain, the mutants (m-7 and m-56) showed higher inhibited ratio on the plates adding 1.0 and 2.0 mM SA, and their growth was completely inhibited by 4.0 mM SA. However, the complementary strains restored deficiencies of the mutants and exhibited similar

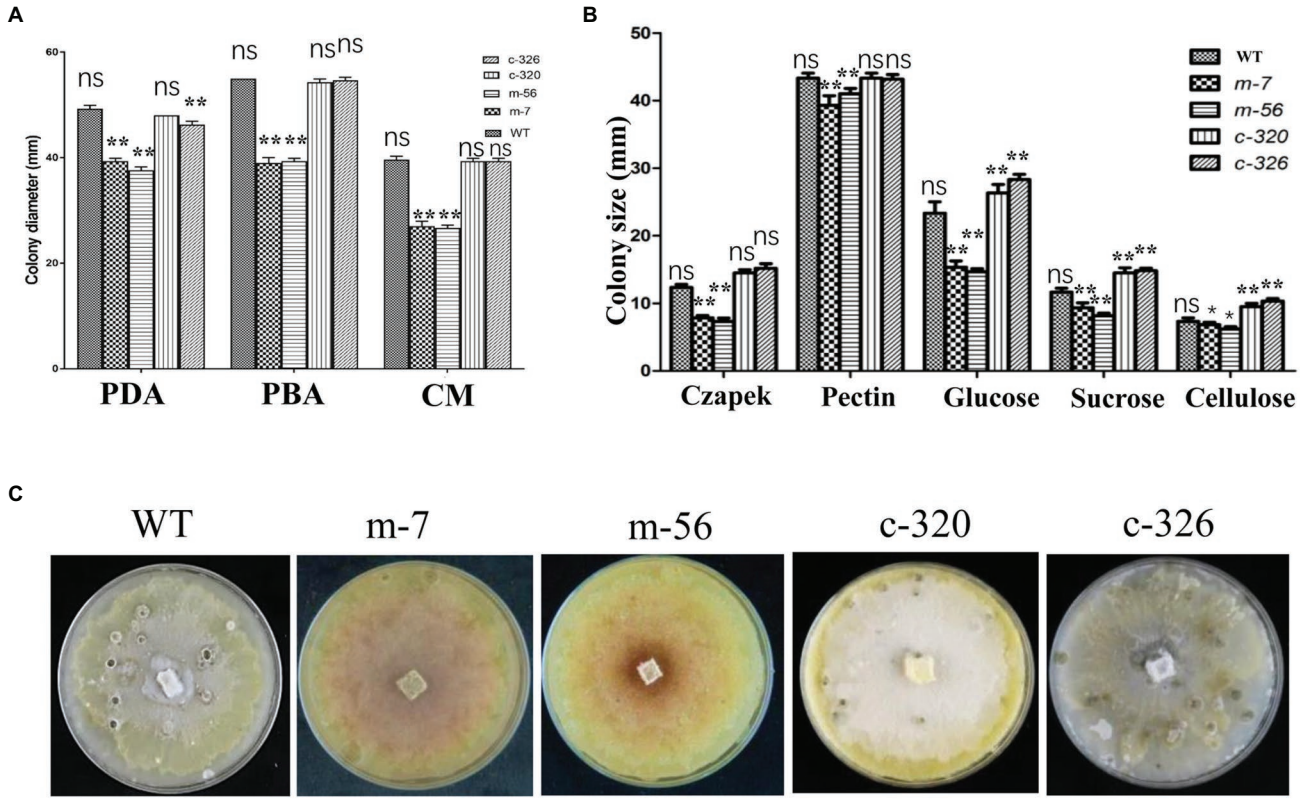


FIGURE 2 | Characterization of growth and development of the *vpxlnR* mutants. **(A)** Colony growth on potato dextrose agar (PDA), pear branch agar (PBA), and CM media. Mycelial agar plugs of wild-type (WT), *vpxlnR mutants* (m-7 and m-56), and *VpxlnR* complementation (C-320 and C-326) were placed on PDA, PBA, and CM media. Colony diameter was measured at 36 h, and images were captured at the same time. **(B)** Radial growth of WT, *vpxlnR mutants*, and complementation on carbon sources. Colony diameters were measured at 36 h. Czapek's medium supplemented with glucose, sucrose, pectin, and cellulose, and Czapek only served as the control. The plates were incubated at 25°C in the dark for 48 h. Typical images of the colonies were captured at 36 h. **(C)** Fruiting body formation. Fruiting bodies were induced under 16-h light/8-h dark for 20 days on PDA medium. Each strain was cultured on three agar plates. Images were captured at 20 days. In the image, ns indicates a value of $p > 0.05$, * indicates a value of $p < 0.05$, ** indicates a value of $p < 0.01$, and the statistical analysis was performed using two-way ANOVA.

or lower inhibited ratio with increasing SA concentration (**Figure 4A**). These results indicate that VpxlnR has a large impact on SA stress. Similar to the SA stress assays, the mutants exhibited increased susceptibility to H_2O_2. With an increasing concentration of H_2O_2, the inhibited ratio of the mutants were greatly increased compared with those of the wild-type and complementary strains (**Figure 4B**). When the H_2O_2 concentration was 2 mM, mycelial growth of the m-56 mutant was completely inhibited. Similarly, the growth of the m-7 mutant was also significantly reduced, and its colony became abnormal, showing a thinner layer than the wild-type and the complementary strains (**Figure 4B**). These results indicate that VpxlnR is involved in H_2O_2 stress in *V. pyri*. Taken together, these results demonstrate that VpxlnR functions in response to overcoming host immunity to aid infection.

Prediction of VpxlnR Binding Promoters

To characterize genes controlled by VpxlnR, we obtained candidate promoter sequences (2000 bp upstream region from the initial codon of each gene) by searching CGG triplets or TTAGSCTAA in *V. pyri*. As a result, we found 354 promoters that have similar motifs, among which 268 promoters were present on the antisense strand, and 86 promoters were present on the sense strand (**Supplementary Table S1**). The genes downstream of each promoter were also obtained using Seqhunter 1.0, and the COG functional annotations were analyzed. A total of 148 genes had COG terms, and the most enriched terms were involved in "carbohydrate transport and metabolism," "lipid transport and metabolism," "secondary metabolite biosynthesis, transport, and catabolism," and "general function prediction only" (**Figure 5**; **Supplementary Table S2**), which suggests that VpxlnR controls carbon resource utilization. However, there is limited information on the genes involved in virulence. To explore genes involved in virulence, we selected all 354 genes and performed PHI-BLAST. We obtained 28 proteins from the PHI-BLAST hits. Thirteen of them play an important role in virulence in other fungi because the corresponding mutants showed reduced virulence, and 15 were annotated as fatal (**Supplementary Table S3**). These data indicate that these genes may contribute to virulence downstream of VpxlnR in *V. pyri*.

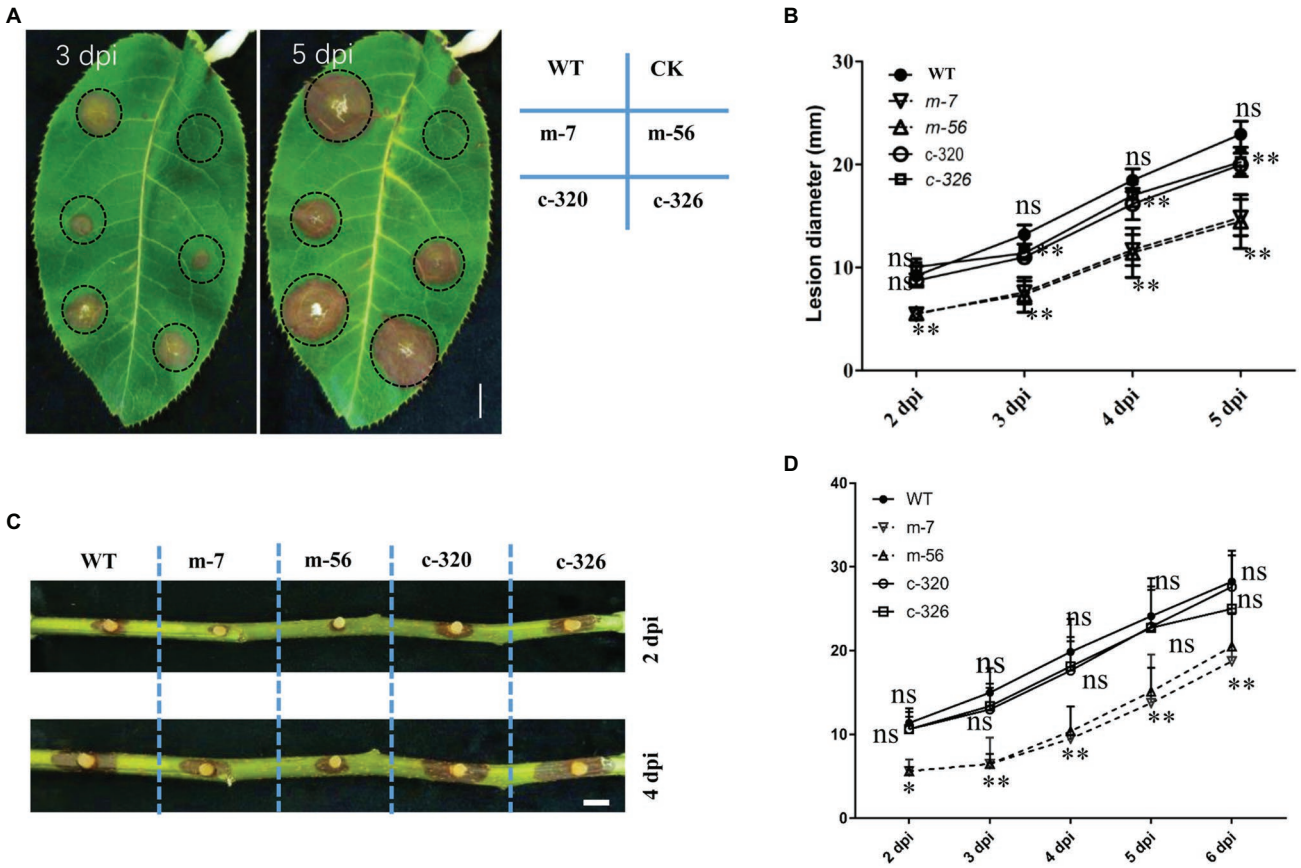


FIGURE 3 | Effect of *VpxlnR* on virulence. **(A)** Virulence of the *vpxlnR* mutants on pear leaves. Mycelial agar plugs of the strains, including WT, *vpxlnR* mutants, and complementation, were inoculated on pear leaves. Photos were taken at 3 and 5 dpi, respectively. Lesion diameters were measured at 2–5 dpi. **(B)** Lesion diameters on pear leaves. Lesion diameters were collected on eight leaves. **(C)** Pathogenicity of the *vpxlnR* mutants on pear branches. Related strains were inoculated on wounded 1-year-old branches, and branches were incubated at 25°C in the dark. Branch images were captured at 2 and 4 dpi. Lesion diameters were measured at 2–6 dpi. **(D)** Lesion lengths on pear branches. Lesion lengths were obtained from 13 inoculated branches. Bar = 1 cm, ns indicates a value of $p > 0.05$, * indicates a value of $p < 0.05$, and ** indicates a value of $p < 0.01$; the statistical analysis was performed by two-way ANOVA.

Identification of Virulence- or Growth-Related Genes Binding to VpxlnR

According to gene annotation and PHI-BLAST analysis, we chose the promoters of the eight genes encoding salicylate hydroxylase (VP1G_09520), alpha-xylosidase (VP1G_06369), G-protein beta subunit (VP1G_02856), acid phosphatase (VP1G_03782), putative phosphotransferases (VP1G_03516), and three serine/threonine-protein kinases (VP1G_03128, VP1G_04075, and VP1G_10966) for further characterization. To confirm whether VpxlnR binds to the promoter regions of virulence- or growth-related genes *in vivo*, we performed Y1H assays. The Y1H results showed that VpxlnR could bind to the promoter regions of the genes encoding salicylate hydroxylase (VP1G_09520), serine/threonine-protein kinase (VP1G_03128), alpha-xylosidase (VP1G_06369), G-protein beta subunit (VP1G_02856), and acid phosphatase (VP1G_03782; **Figure 6**), suggesting that VpxlnR may determine virulence or growth by controlling the expression of extracellular proteins and signaling pathways. VpxlnR did not bind to the promoters of the other three genes (data not shown). Furthermore, we predicted that the DNA-binding motifs of all three target genes (VP1G_09520, VP1G_03782, and VP1G_02856) contain the degenerate sequence MBSGTCCGY (**Supplementary Table S4**). Another two genes (VP1G_03128 and VP1G_06369) contain a motif including a GGC triplet; however, they were not the same as the reported motif interacting with xlnR orthologues (**Supplementary Table S4**). These results suggest that VpxlnR could interact with motifs consisting of MBSGTCCGY and GGC triplets *in vivo*.

Expression Levels of the Target Genes Controlled by VpxlnR

To assess whether the expression levels of five target genes were affected by *VpxlnR* deletion, we analyzed their transcript levels by RT–qPCR. Based on RT–qPCR results, we found that the expression levels of two genes (VP1G_06369 and VP1G_03128) were only reduced in one mutant (m-7) and not in m-56 (**Figures 7A,B**). These results demonstrated that these two genes (VP1G_06369 and VP1G_03128) might also be controlled by regulators other than xlnR. The other three

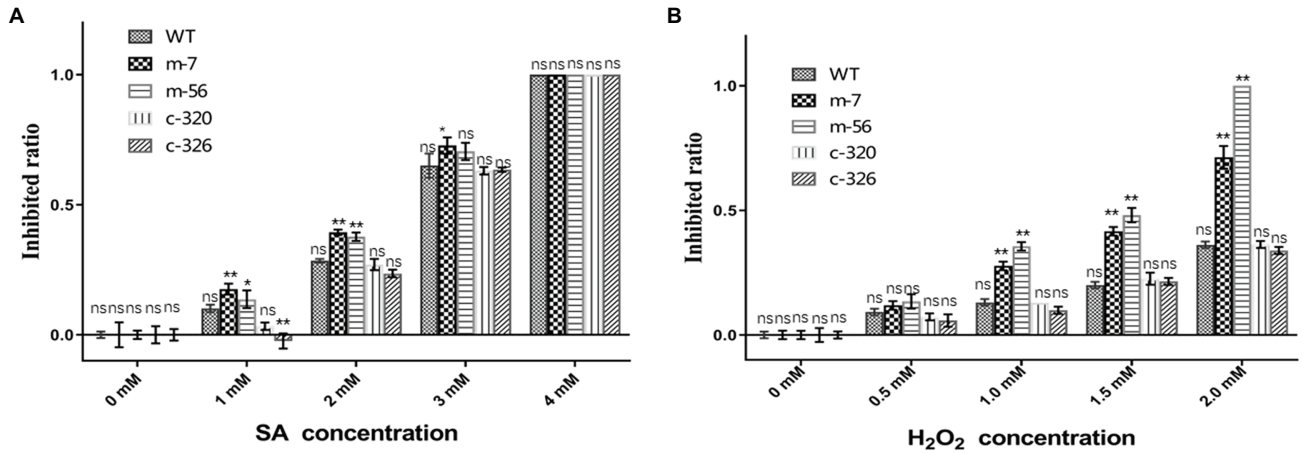


FIGURE 4 | Susceptibility of the *VpxlnR* mutant to host mimic stresses. **(A)** Effect of SA stress on strain growth. Wild-type (WT), VpxlnR deletion mutant (m-7 and m-56), and complementation (C-320 and C-326) strains were grown on PDA medium supplemented with the indicated SA concentrations and incubated at 25°C for 36 h. **(B)** Impact of oxidative stress on colony growth. Strains were cultured on PDA medium supplemented with H_2O_2 and incubated at 25°C for 36 h. Each treatment was repeated three times, and colony diameters were recorded at 36 h. The data were recorded at 36 h. The statistical analysis was performed by two-way ANOVA, ns indicates a value of $p > 0.05$, * indicates a value of $p < 0.05$, and ** indicates a value of $p < 0.01$.

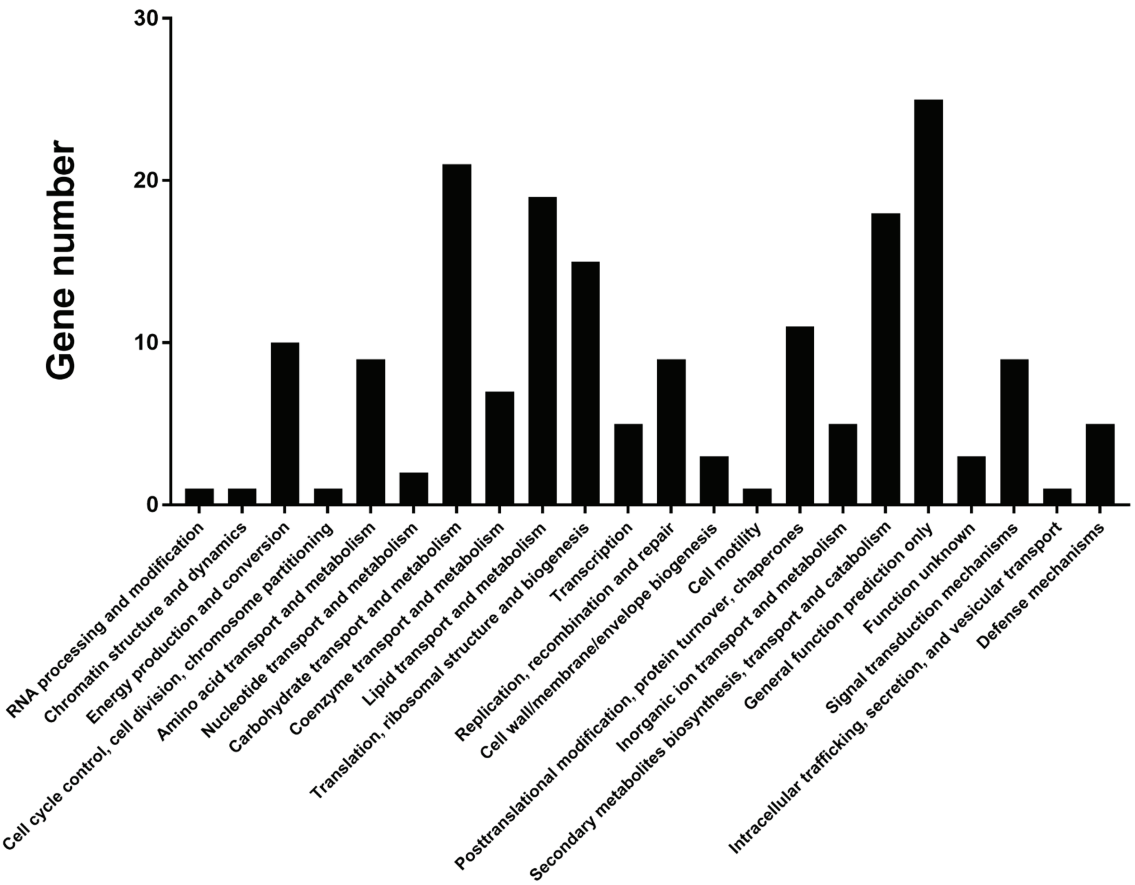


FIGURE 5 | COG annotation of the predicted target genes. Approximately 148 target genes have COG terms. COG terms are indicated on the image. *X*-axis: COG classes. *Y*-axis: shown in the image.

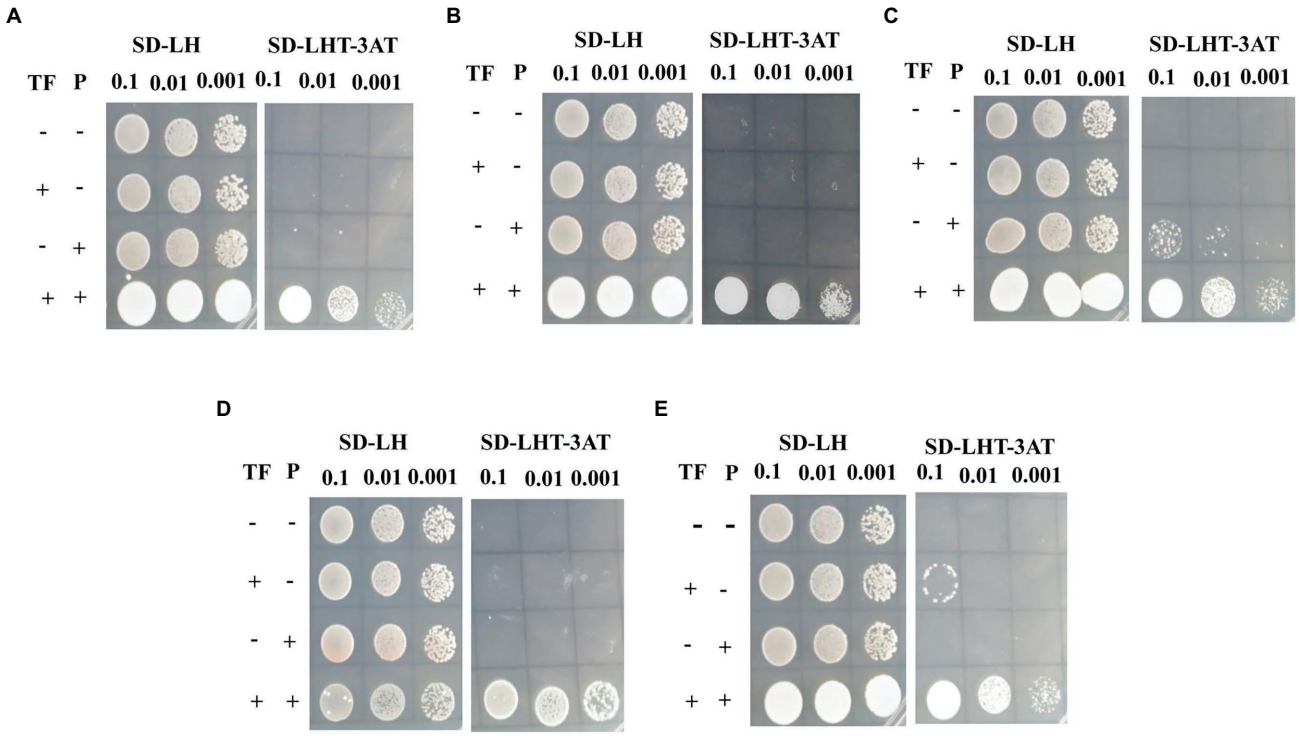


FIGURE 6 | Assessment of *VpxlnR* binding activity using yeast one-hybrid (Y1H). Genes were selected from PHI-BLAST analysis, including **(A)** VP1G_09520: salicylate hydroxylase, **(B)** VP1G_03128: serine/threonine-protein kinase, **(C)** VP1G_06369: alpha-xylosidase, **(D)** VP1G_02856: G-protein beta subunit, and **(E)**. VP1G_03782: acid phosphatase. *VpxlnR* and pHis2::P coexpressed in yeast cells were cultured on SD/−Leu/His plates, and positive clones were screened on SD/−Trp/−Leu/-His plates with 3AT plates. TF: VpxlnR, P: promoter. Empty pGADT7 (–) and empty pHIS2 (–) plasmids, and pGADT7::*VpxlnR*/empty pHIS2 vector and empty pGADT7/pHIS2::P were used as controls. (+) indicates that the yeast cells expressed *VpxlnR or* contained gene promoters. The liquid cultures (OD_{600} = 0.5) were diluted as shown in the image. Positive yeast clones were formed after culture on SD/−Trp/−Leu/-His/3AT for 3 days, and images were captured.

genes were downregulated in the two mutants (m-7 and m-56; **Figures 7C–E**), suggesting that VpxlnR could regulate the expression levels of these genes by directly binding gene promoters. These results demonstrated that VpxlnR could positively control the transcript levels of virulence-related genes to contribute to virulence in *V. pyri*.

DISCUSSION

Valsa pyri is a woody pathogen causing trunk canker disease on pear trees in most orchards of northern China. The fungus can infect pear or apple trees, leaves, and fruits, resulting in great yield loss or tree death (Abe et al., 2007; Xu et al., 2018). *Valsa pyri* is a necrotrophic fungus that can penetrate the phloem and xylem. Transcriptomic analysis also showed that cellulose- or pectin-degrading enzyme-encoding genes were significantly upregulated in the infection stage (He et al., 2018). However, cellulase and pectinase were not the key factors determining virulence in *V. mali* (Yin et al., 2015). Previous studies showed that several proteins secreted by *V. mali* function as effectors, which could cause cell death or change the immune system in *Nicotiana benthamiana* (Li et al., 2015b; Feng et al., 2018; Zhang et al., 2019). However, the deletions of these protein-encoding genes in *V. mali* did not lead to a great virulence reduction. Moreover, woody plants are obviously distinguished from herbs, especially trunk diseases. Therefore, studying woody plant pathogens faces greater challenges. VpxlnR and its orthologues generally contain the GAL4 domain and have important roles in controlling xylanolytic activity (van Peij et al., 1998), degradation of the polysaccharides xylan and cellulose (Hasper et al., 2000), and virulence (Calero-Nieto et al., 2007). Due to its potential function in the degradation of xylan and cellulose, which are key chemical components in the trunk or branches of pear trees, VpxlnR may contribute greatly to virulence in *V. pyri*.

The VpxlnR gene contains four introns, and the genome sequence is not well assembled at its location (Yin et al., 2015); thus, the transcriptome *de novo* analysis results provide more clues to identify this gene (**Supplementary Figure S1**). The xlnR protein predicted in the *V. pyri* genome lost a GAL4 domain. However, our study showed that it contains a GAL4 domain. Additionally, most xlnR orthologues represent one zinc cluster, especially VmxlnR, which shares high identity with VpxlnR. We deduced that xlnR proteins generally contains the GAL4 domain, and those orthologues with the GAL4 domain lost may because of the wrong prediction. Therefore,

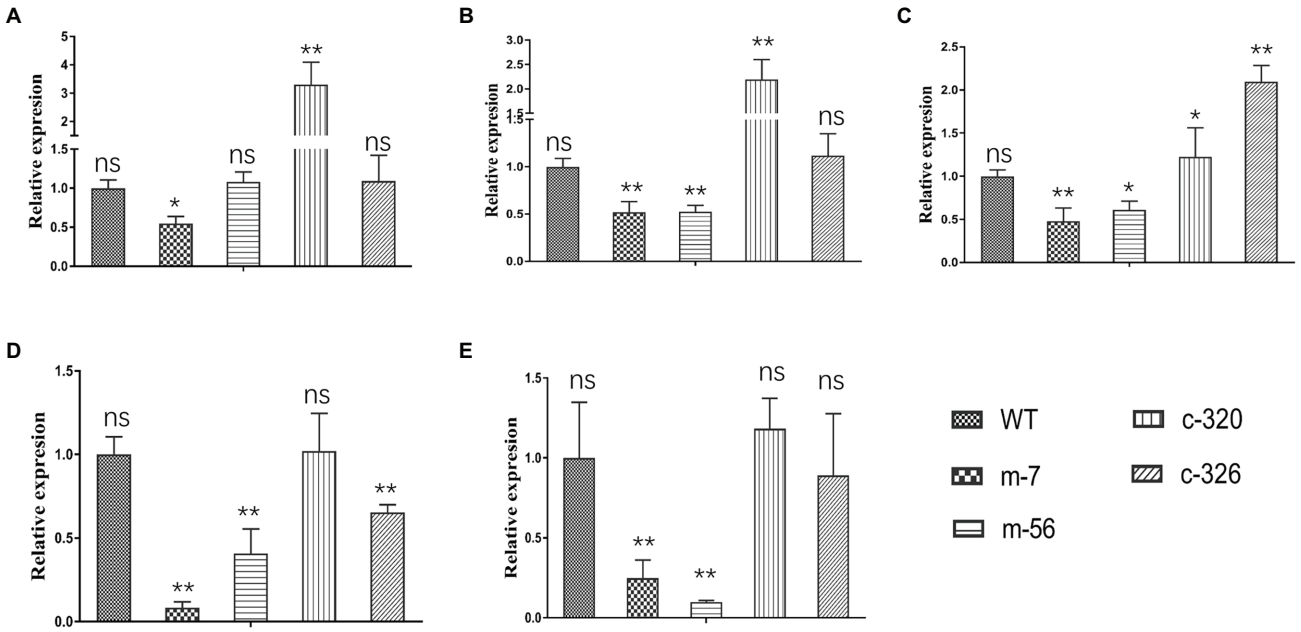


FIGURE 7 | RT–qPCR test of target genes controlled by VpxlnR. **(A)** Salicylate hydroxylase encoding gene (VP1G_09520), **(B)** serine/threonine-protein kinase encoding gene (VP1G_03128), **(C)** alpha-xylosidase encoding gene (VP1G_06369), **(D)** G-protein beta subunit encoding gene (VP1G_02856), and **(E)** Acid phosphatase encoding gene (VP1G_03782). The transcript level of each gene was normalized to actin expression. Each test was repeated three times, and the statistical analysis was performed using one-way ANOVA. ns indicates a value of $p > 0.05$, * indicates a value of $p < 0.05$, and ** indicates a value of $p < 0.01$.

we think predictions for xlnR orthologues should be rigorous, and the genomes are not sufficient to identify this gene in other fungus.

The main functions of xlnR orthologues in *Aspergillus* spp. are to regulate xylose and other polysaccharides (van Peij et al., 1998; Hasper et al., 2004; Noguchi et al., 2011; Ishikawa et al., 2018; Khosravi et al., 2019). However, AnxlnR was not involved in glucose or sucrose utilization. This result implies that xlnR is involved in many aspects of fungal life. In our study, we found that the growth ratio of the *VpxlnR* deletion mutants was significantly increased on glucose compared with the no-glucose control, but compared with the wild-type, their growth ratio was greatly reduced. Thus, VpxlnR should be a key regulator controlling growth by acquiring glucose from the environment or host. When sucrose was added, the results were similar to those of the control, which indicates that VpxlnR does not participate in sucrose utilization. Similar to a previous study (de Souza et al., 2013; Kowalczyk et al., 2015; Llanos et al., 2019), the colony size of the mutants did not decrease significantly on pectin or cellulose plates compared with Vp297, which implies that the mutants may grow normally in the infection stage because of their abilities to utilize cell wall components such as pectin and cellulose. These outcomes are partially in agreement with previous studies and provide new insights into the function of xlnR orthologues.

Previously, FoxlnR did not affect the virulence of *F. oxysporum* on tomato fruits, but it actually affected the pathogenicity of *F. oxysporum* on tomato roots before 10 dpi (Calero-Nieto et al., 2007). It mainly controls the transcript levels of xylanase-encoding genes (Calero-Nieto et al., 2007). In our study, deficiency of the *VpxlnR* deletion mutants on infection was detected on pear branches or leaves (**Figure 3**). Moreover, the *VpxlnR* deletion mutants were also sensitive to H_2O_2 stress and SA stress in a concentration-dependent manner (**Figure 3**). Interestingly, the Δ*xlnR* strains of *A. niger* were sensitive to oxidative stress when grown on media supplemented with glucose (Raulo et al., 2016). Additionally, the mutants reduced their abilities to utilize glucose or sucrose (**Figure 2B**). We believe that host extracellular glucose or sucrose might induce intracellular ROS levels in fungi, which leads to growth limitation of the *xlnR* mutants. In the early infection stage, the extracellular space of the host cell contains glucose, sucrose, and host defense chemicals such as ROS and SA, which may impair infectious growth of the mutants in host tissue. However, when they overcome innate immunity, they can break the host cell wall and use pectin, cellulose, and xylose and then exhibit restored virulence. In brief, the *VpxlnR* deletion mutants exhibit susceptibility to the host defense response, resulting in reduced virulence at the infection stage of *V. pyri*.

Zn2Cys6 TFs control target genes by binding a motif including a CGG triplet (Cho et al., 2013; Luo et al., 2016), while XlnR interacts with promoters containing not only CGG triplets but also TTAGSCTAA motifs (Ishikawa et al., 2018). To broadly explore the target genes of VpxlnR, we obtained all of the promoters containing NRHKGMCCGM in the *V. pyri* genome (Cho et al., 2013; Ishikawa et al., 2018). Eight promoters more than 1,500 bp in length were identified and five of them could interact with VpxlnR in Y1H assays. These results suggest that

the motif consisting of MDSGTCCGY is very likely to interact with VpxlnR. Moreover, VpxlnR also recognizes a motif containing GGC triplets, although no specific feature was observed in the 3′ flanking sequence of the GGC triplet. These results extend the binding site of xlnR and provide new clues for studying zinc cluster-type TFs.

Genes directly controlled by VpxlnR contain secreted proteins such as alpha-xylosidase and acid phosphatase, signaling transduction regulators including G-protein beta subunit and serine/threonine-protein kinase, and salicylate hydroxylase (**Figure 6**). There were also three genes encoding putative serine/threonine-protein kinase KIN28 (VP1G_04075) and GCN2 (VP1G_10966) and putative phosphotransferases (VP1G_03516) that could not interact with VpxlnR, but their transcript levels were significantly reduced in the deletion mutants (**Supplementary Figure S4**). These results suggest that they might function downstream of VpxlnR in regulating virulence. However, there are very limited reports on the function of these two secreted proteins in fungal virulence. G-protein and its regulators could regulate conidiation, antioxidant capacity, and virulence in *M. robertsii*, *Mucor circinelloides*, *Cryphonectria parasitica*, and *Ustilago maydis* (Segers et al., 2004; Moretti et al., 2017; Tong et al., 2020; Valle-Maldonado et al., 2020). Moreover, the serine/threonine-protein kinase ChSch9 participated in the virulence of *Colletotrichum higginsianum* (Sohail et al., 2021). Therefore, we deduced that VpxlnR responds to H_2O_2 stress, virulence, and fruiting body formation by controlling the G-protein and serine/threonine-protein kinase encoding genes. Additionally, whether the two secreted proteins contribute to the virulence of *V. pyri* should be confirmed in the future.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

FH designed the experiments, analyzed the data, and wrote the manuscript. A-MK, FH, and JY finished all most experiments. YZ and FL provided funding for this study. ZF revised the manuscript. The others participated in experiments or manuscript revision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the Anhui Provincial Natural Science Foundation (2008085MC77 and 2008085QC129), the Program of Fujian Key Laboratory for Monitoring and Management of Crop Pests (026052020008), the Major Science and Technology Projects in Anhui Province (No. 2020003a06020009), and the Earmarked Fund for China Agriculture Research System (CARS-28-16).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.784686/full#supplementary-material

Supplementary Figure S1 | Distinct sites of the sequences predicted in transcriptomics and genomics of *V. pyri*. VP1G_03432 was characterized using *V. pyri* genomics. The sequence of VpxlnR was obtained from *de novo* transcriptomics. The line shown in the figure indicates the gaps of the two sequences.

Supplementary Figure S2 | Constructs for VpxlnR replacement and complementation. VpxlnR deletion constructs were generated by triple joint PCR amplification (described in the "Materials and Methods" section). The arrows indicate primer sites. The numbers (1–12) represent the primers (**Supplementary Table S5**) used for deletion construct generation and genomic PCR identification.

Supplementary Figure S3 | The mutant was confirmed by genomic PCR and RT–qPCR. **(A)** Genomic PCR identification for fragments in replacement sites. Primer pairs with numbers are shown in the right panel, and numbers are the same as in **Supplementary Figure S1**. WT, wild-type strain (Vp297); m-7 and m-56, VpxlnR deletion mutants; C-320 and C-236, VpxlnR complementation strains. **(B)** VpxlnR expression level in deletion mutant. Transcript levels were quantified by RT–qPCR and normalized to *actin* gene expression. Each test was repeat three times. The relative transcript level of the gene at time point 0 was set as 1.0. The data were analyzed using one-way ANOVA. ns indicates a value of $p > 0.05$, ** indicates a value of $p < 0.01$.

Supplementary Figure S4 | RT–qPCR test of nontarget genes controlled by VpxlnR. **(A)** Serine/threonine-protein kinase KIN28-encoding gene (VP1G_04075); **(B)** serine/threonine-protein kinase GCN2-encoding gene (VP1G_10966); **(C)** putative phosphotransferase-encoding gene (VP1G_03516). The transcript level of each gene was normalized to actin expression. Each test was repeated three times, and the data were analysed using one-way ANOVA. ns indicates a value of $p > 0.05$, * indicates a value of $p < 0.05$, and ** indicates value of $p < 0.01$.

REFERENCES

Abe, K., Kotoda, N., Kato, H., and Soejima, J. (2007). Resistance sources to Valsa canker (*Valsa ceratosperma*) in a germplasm collection of diverse Malus species. *Plant Breed.* 126, 449–453. doi: 10.1111/j.1439-0523.2007.01379.x

Aro, N., Saloheimo, A., Ilmen, M., and Penttila, M. (2001). ACEII, a novel transcriptional activator involved in regulation of cellulase and xylanase genes of *Trichoderma reesei*. *J. Biol. Chem.* 276, 24309–24314. doi: 10.1074/jbc.M003624200

Battaglia, E., Klaubauf, S., Vallet, J., Ribot, C., Lebrun, M. H., and de Vries, R. P. (2013). Xlr1 is involved in the transcriptional control of the pentose catabolic pathway, but not hemi-cellulolytic enzymes in *Magnaporthe oryzae*. *Fungal Genet. Biol.* 57, 76–84. doi: 10.1016/j.fgb.2013.06.005

Calero-Nieto, F., Di Pietro, A., Roncero, M. I., and Hera, C. (2007). Role of the transcriptional activator xlnR of Fusarium oxysporum in regulation of xylanase genes and virulence. *Mol. Plant-Microbe Interact.* 20, 977–985. doi: 10.1094/MPMI-20-8-0977

Cho, Y., Ohm, R. A., Grigoriev, I. V., and Srivastava, A. (2013). Fungal-specific transcription factor AbPf2 activates pathogenicity in *Alternaria brassicicola*. *Plant J.* 75, 498–514. doi: 10.1111/tpj.12217

Choi, J., Kim, Y., Kim, S., Park, J., and Lee, Y. H. (2009). MoCRZ1, a gene encoding a calcineurin-responsive transcription factor, regulates fungal growth

and pathogenicity of *Magnaporthe oryzae*. *Fungal Genet. Biol.* 46, 243–254. doi: 10.1016/j.fgb.2008.11.010

de Souza, W. R., Maitan-Alfenas, G. P., de Gouvea, P. F., Brown, N. A., Savoldi, M., Battaglia, E., et al. (2013). The influence of *Aspergillus niger* transcription factors AraR and XlnR in the gene expression during growth in D-xylose, L-arabinose and steam-exploded sugarcane bagasse. *Fungal Genet. Biol.* 60, 29–45. doi: 10.1016/j.fgb.2013.07.007

de Vries, R. P., Visser, J., and de Graaff, L. H. (1999). CreA modulates the XlnR-induced expression on xylose of *Aspergillus niger* genes involved in xylan degradation. *Res. Microbiol.* 150, 281–285. doi: 10.1016/S0923-2508(99)80053-9

Feng, Y., Yin, Z., Wu, Y., Xu, L., Du, H., Wang, N., et al. (2020). LaeA controls virulence and secondary metabolism in apple canker pathogen *Valsa mali*. *Front. Microbiol.* 11:581203. doi: 10.3389/fmicb.2020.581203

Feng, H., Zhang, M., Zhao, Y., Li, C., Song, L., and Huang, L. (2018). Secreted peroxidases VmPODs play critical roles in the conidiation, H2O2 sensitivity and pathogenicity of *Valsa mali*. *Fungal Genet. Biol.* 119, 20–28. doi: 10.1016/j.fgb.2018.08.003

Fujii, T., Inoue, H., and Ishikawa, K. (2014). Characterization of the xylanase regulator protein gene, xlnR, in *Talaromyces cellulolyticus* (formerly known as *Acremonium cellulolyticus*). *Biosci. Biotechnol. Biochem.* 78, 1564–1567. doi: 10.1080/09168451.2014.923298

Guo, M., Chen, Y., Du, Y., Dong, Y. H., Guo, W., Zhai, S., et al. (2011). The bZIP transcription factor MoAP1 mediates the oxidative stress response and is critical for pathogenicity of the Rice blast fungus *Magnaporthe oryzae*. *PLoS Pathog.* 7:e1001302. doi: 10.1371/journal.ppat.1001302

Hasper, A. A., Trindade, L. M., van der Veen, D., van Ooyen, A. J. J., and de Graaff, L. H. (2004). Functional analysis of the transcriptional activator XlnR from *Aspergillus niger*. *Microbiology* 150, 1367–1375. doi: 10.1099/mic.0.26557-0

Hasper, A. A., Visser, J., and de Graaff, L. H. (2000). The Aspergillus Niger transcriptional activator XlnR, which is involved in the degradation of the polysaccharides xylan and cellulose, also regulates D-xylose reductase gene expression. *Mol. Microbiol.* 36, 193–200. doi: 10.1046/j.1365-2958.2000.01843.x

He, F., Zhang, X., Li, B. X., Safdar, A., Ai, G., Kange, A. M., et al. (2018). Comparative transcriptomics of two *Valsa pyri* isolates uncover different strategies for virulence and growth. *Microb. Pathog.* 123, 478–486. doi: 10.1016/j.micpath.2018.08.013

He, F., Zhang, X., Mafurah, J. J., Zhang, M. X., Qian, G. L., Wang, R. B., et al. (2016). The transcription factor VpCRZ1 is required for fruiting body formation and pathogenicity in *Valsa pyri*. *Microb. Pathog.* 95, 101–110. doi: 10.1016/j.micpath.2016.02.018

Ishikawa, K., Kunitake, E., Kawase, T., Atsumi, M., Noguchi, Y., Ishikawa, S., et al. (2018). Comparison of the paralogous transcription factors AraR and XlnR in *Aspergillus oryzae*. *Curr. Genet.* 64, 1245–1260. doi: 10.1007/s00294-018-0837-5

Kange, A. M., Xia, A., Si, J., Li, B., Zhang, X., Ai, G., et al. (2019). The fungal-specific transcription factor VpFSTF1 is required for virulence in *Valsa pyri*. *Front. Microbiol.* 10:2945. doi: 10.3389/fmicb.2019.02945

Kange, A. M., Xia, A., Si, J. R., Li, B. X., Zhang, X., Ai, G., et al. (2020). The fungal-specific transcription factor VpFSTF1 is required for virulence in *Valsa pyri*. *Front. Microbiol.* 10:2945. doi: 10.3389/fmicb.2019.02945

Katz, M. E., Braunberger, K., Yi, G., Cooper, S., Nonhebel, H. M., and Gondro, C. (2013). A p53-like transcription factor similar to Ndt80 controls the response to nutrient stress in the filamentous fungus, *Aspergillus nidulans* 2:72. doi: 10.12688/f1000research.2-72.v1

Khosravi, C., Kowalczyk, J. E., Chroumpi, T., Battaglia, E., Aguilar Pontes, M. V., Peng, M., et al. (2019). Transcriptome analysis of Aspergillus Niger xlnR and xkiA mutants grown on corn Stover and soybean hulls reveals a highly complex regulatory network. *BMC Genomics* 20:853. doi: 10.1186/s12864-019-6235-7

Kim, S., Park, S. Y., Kim, K. S., Rho, H. S., Chi, M. H., Choi, J., et al. (2009). Homeobox transcription factors are required for conidiation and appressorium development in the rice blast fungus Magnaporthe oryzae. *PLoS Genet.* 5:e1000757. doi: 10.1371/journal.pgen.1000757

Klaubauf, S., Narang, H. M., Post, H., Zhou, M., Brunner, K., Mach-Aigner, A. R., et al. (2014). Similar is not the same: differences in the function of the (hemi-)cellulolytic regulator XlnR (Xlr1/Xyr1) in filamentous fungi. *Fungal Genet. Biol.* 72, 73–81. doi: 10.1016/j.fgb.2014.07.007

Kowalczyk, J. E., Gruben, B. S., Battaglia, E., Wiebenga, A., Majoor, E., and de Vries, R. P. (2015). Genetic interaction of *Aspergillus nidulans* galR, xlnR and araR in regulating D-Galactose and L-arabinose release and catabolism gene expression. *PLoS One* 10:e0143200. doi: 10.1371/journal.pone.0143200

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Li, C., Melesse, M., Zhang, S., Hao, C., Wang, C., Zhang, H., et al. (2015a). FgCDC14 regulates cytokinesis, morphogenesis, and pathogenesis in *Fusarium graminearum*. *Mol. Microbiol.* 98, 770–786. doi: 10.1111/mmi.13157

Li, Z. P., Yin, Z. Y., Fan, Y. Y., Xu, M., Kang, Z. S., and Huang, L. L. (2015b). Candidate effector proteins of the necrotrophic apple canker pathogen *Valsa mali* can suppress BAX-induced PCD. *Front. Plant Sci.* 6:579. doi: 10.3389/fpls.2015.00579

Llanos, A., Dejean, S., Neugnot-Roux, V., Francois, J. M., and Parrou, J. L. (2019). Carbon sources and XlnR-dependent transcriptional landscape of CAZymes in the industrial fungus Talaromyces versatilis: when exception seems to be the rule. *Microb. Cell Factories* 18:14. doi: 10.1186/s12934-019-1062-8

Lu, J., Cao, H., Zhang, L., Huang, P., and Lin, F. (2014). Systematic analysis of Zn2Cys6 transcription factors required for development and pathogenicity by high-throughput gene knockout in the rice blast fungus. *PLoS Pathog.* 10:e1004432. doi: 10.1371/journal.ppat.1004432

Luo, X., Mao, H., Wei, Y., Cai, J., Xie, C., Sui, A., et al. (2016). The fungal-specific transcription factor Vdpf influences conidia production, melanized microsclerotia formation and pathogenicity in *Verticillium dahliae*. *Mol. Plant Pathol.* 17, 1364–1381. doi: 10.1111/mpp.12367

MacPherson, S., Larochelle, M., and Turcotte, B. (2006). A fungal family of transcriptional regulators: the zinc cluster proteins. *Microbiol. Mol. Biol. Rev.* 70, 583–604. doi: 10.1128/MMBR.00015-06

Marui, J., Tanaka, A., Mimura, S., de Graaff, L. H., Visser, J., Kitamoto, N., et al. (2002). A transcriptional activator, AoXlnR, controls the expression of genes encoding xylanolytic enzymes in *Aspergillus oryzae*. *Fungal Genet. Biol.* 35, 157–169. doi: 10.1006/fgbi.2001.1321

Moretti, M., Wang, L., Grognet, P., Lanver, D., Link, H., and Kahmann, R. (2017). Three regulators of G protein signaling differentially affect mating, morphology and virulence in the smut fungus *Ustilago maydis*. *Mol. Microbiol.* 105, 901–921. doi: 10.1111/mmi.13745

Noguchi, Y., Tanaka, H., Kanamaru, K., Kato, M., and Kobayashi, T. (2011). Xylose triggers reversible phosphorylation of XlnR, the fungal transcriptional activator of Xylanolytic and cellulolytic genes in *Aspergillus oryzae*. *Biosci. Biotech. Bioch.* 75, 953–959. doi: 10.1271/bbb.100923

Oka, H., Kojima, T., Ihara, K., Kobayashi, T., and Nakano, H. (2019). Comprehensive investigation of the gene expression system regulated by an *Aspergillus oryzae* transcription factor XlnR using integrated mining of gSELEX-Seq and microarray data. *BMC Genomics* 20:16. doi: 10.1186/s12864-018-5375-5

Raulo, R., Kokolski, M., and Archer, D. B. (2016). The roles of the zinc finger transcription factors XlnR, ClrA and ClrB in the breakdown of lignocellulose by *Aspergillus niger*. *AMB Express* 6:5. doi: 10.1186/s13568-016-0177-0

Rauscher, R., Wurleitner, E., Wacenovsky, C., Aro, N., Stricker, A. R., Zeilinger, S., et al. (2006). Transcriptional regulation of xyn1, encoding xylanase I, in *Hypocrea jecorina*. *Eukaryot. Cell* 5, 447–456. doi: 10.1128/EC.5.3.447-456.2006

Rybak, K., See, P. T., Phan, H. T., Syme, R. A., Moffat, C. S., Oliver, R. P., et al. (2017). A functionally conserved Zn2 Cys6 binuclear cluster transcription factor class regulates necrotrophic effector gene expression and host-specific virulence of two major *Pleosporales* fungal pathogens of wheat. *Mol. Plant Pathol.* 18, 420–434. doi: 10.1111/mpp.12511

Segers, G. C., Regier, J. C., and Nuss, D. L. (2004). Evidence for a role of the regulator of G-protein signaling protein CPRGS-1 in Galpha subunit CPG-1-mediated regulation of fungal virulence, conidiation, and hydrophobin synthesis in the chestnut blight fungus *Cryphonectria parasitica*. *Eukaryot. Cell* 3, 1454–1463. doi: 10.1128/EC.3.6.1454-1463.2004

Shelest, E. (2017). Transcription factors in fungi: TFome dynamics, three major families, and dual-specificity TFs. *Front. Genet.* 8:53. doi: 10.3389/fgene.2017.00053

Sohail, M. A., Yuan, Q. F., Yan, Y. Q., Huang, J. B., Hsiang, T., and Zheng, L. (2021). ChSch9 is required for infection-related morphogenesis and pathogenicity in *Colletotrichum higginsianum*. *Can. J. Plant Pathol.* 43, 871–885. doi: 10.1080/07060661.2021.1921850

Sun, J., Tian, C., Diamond, S., and Glass, N. L. (2012). Deciphering transcriptional regulatory mechanisms associated with hemicellulose degradation in *Neurospora crassa*. *Eukaryot. Cell* 11, 482–493. doi: 10.1128/EC.05327-11

Tamayo, E. N., Villanueva, A., Hasper, A. A., de Graaff, L. H., Ramon, D., and Orejas, M. (2008). CreA mediates repression of the regulatory gene xlnR which controls the production of xylanolytic enzymes in *Aspergillus nidulans*. *Fungal Genet. Biol.* 45, 984–993. doi: 10.1016/j.fgb.2008.03.002

Tong, Y., Wu, H., Liu, Z., Wang, Z., and Huang, B. (2020). G-protein subunit Galphai in mitochondria, MrGPA1, affects Conidiation, stress resistance, and virulence of Entomopathogenic fungus *Metarhizium robertsii*. *Front. Microbiol.* 11:1251. doi: 10.3389/fmicb.2020.01251

Umesha, S., Manukumar, H. M., and Raghava, S. (2016). A rapid method for isolation of genomic DNA from food-borne fungal pathogens. *3 Biotech* 6:123. doi: 10.1007/s13205-016-0436-4

Valle-Maldonado, M. I., Patino-Medina, J. A., Perez-Arques, C., Reyes-Mares, N. Y., Jacome-Galarza, I. E., Ortiz-Alvarado, R., et al. (2020). The heterotrimeric G-protein beta subunit Gpb1 controls hyphal growth under low oxygen conditions through the protein kinase A pathway and is essential for virulence in the fungus Mucor circinelloides. *Cell. Microbiol.* 22:e13236. doi: 10.1111/cmi.13236

van Peij, N. N., Visser, J., and de Graaff, L. H. (1998). Isolation and analysis of xlnR, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. *Mol. Microbiol.* 27, 131–142. doi: 10.1046/j.1365-2958.1998.00666.x

Veri, A. O., Robbins, N., and Cowen, L. E. (2018). Regulation of the heat shock transcription factor Hsf1 in fungi: implications for temperature-dependent virulence traits. *FEMS Yeast Res.* 18:foy041. doi: 10.1093/femsyr/foy041

Wang, X. L., Zang, R., Yin, Z. Y., Kang, Z. S., and Huang, L. L. (2014). Delimiting cryptic pathogen species causing apple Valsa canker with multilocus data. *Ecol. Evol.* 4, 1369–1380. doi: 10.1002/ece3.1030

Wu, Y., Xu, L., Yin, Z., Feng, H., and Huang, L. (2018). Transcription factor VmSeb1 is required for the growth, development, and virulence in *Valsa mali*. *Microb. Pathog.* 123, 132–138. doi: 10.1016/j.micpath.2018.06.043

Xiong, D., Wang, Y., Tian, L., and Tian, C. (2016). MADS-box transcription factor VdMcm1 regulates Conidiation, Microsclerotia formation, pathogenicity, and secondary metabolism of *Verticillium dahliae*. *Front. Microbiol.* 7:1192. doi: 10.3389/fmicb.2016.01192

Xu, M., Gao, X. N., Chen, J. L., Yin, Z. Y., Feng, H., and Huang, L. L. (2018). The feruloyl esterase genes are required for full pathogenicity of the apple tree canker pathogen *Valsa mali*. *Mol. Plant Pathol.* 19, 1353–1363. doi: 10.1111/mpp.12619

Yin, Z., Liu, H., Li, Z., Ke, X., Dou, D., Gao, X., et al. (2015). Genome sequence of *Valsa* canker pathogens uncovers a potential adaptation of colonization of woody bark. *New Phytol.* 208, 1202–1216. doi: 10.1111/nph.13544

Zhang, W. Q., Gui, Y. J., Short, D. P. G., Li, T. G., Zhang, D. D., Zhou, L., et al. (2018). *Verticillium dahliae* transcription factor VdFTF1 regulates the expression of multiple secreted virulence factors and is required for full virulence in cotton. *Mol. Plant Pathol.* 19, 841–857. doi: 10.1111/mpp.12569

Zhang, M., Xie, S., Zhao, Y., Meng, X., Song, L., Feng, H., et al. (2019). Hce2 domain-containing effectors contribute to the full virulence of *Valsa mali* in a redundant manner. *Mol. Plant Pathol.* 20, 843–856. doi: 10.1111/mpp.12796

Zhao, C., Waalwijk, C., de Wit, P. J., van der Lee, T., and Tang, D. (2011). EBR1, a novel Zn(2)Cys(6) transcription factor, affects virulence and apical dominance of the hyphal tip in Fusarium graminearum. *Mol. Plant-Microbe Interact.* 24, 1407–1418. doi: 10.1094/MPMI-06-11-0158

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 16 March 2022
doi: 10.3389/fmicb.2022.806398

Comparative Genome Analysis Across 128 *Phytophthora* Isolates Reveal Species-Specific Microsatellite Distribution and Localized Evolution of Compartmentalized Genomes

Kajal Mandal[1,2†], Subhajeet Dutta[1,2†], Aditya Upadhyay[1], Arijit Panda[3] and Sucheta Tripathy[1,2*]*

[1] Computational Genomics Laboratory, Department of Structural Biology and Bioinformatics, CSIR-Indian Institute of Chemical Biology, Kolkata, India, [2] Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India, [3] Department of Quantitative Health Science, Mayo Clinic, Rochester, MN, United States

OPEN ACCESS

Edited by:
Danyu Shen,
Nanjing Agricultural University, China

Reviewed by:
Jamie McGowan,
Earlham Institute (EI), United Kingdom
Sophie De Vries,
Dalhousie University, Canada

****Correspondence:***
Arijit Panda
arijpanda@gmail.com
Sucheta Tripathy
tsucheta@iicb.res.in;
tsucheta@gmail.com

† These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *31 October 2021*
Accepted: *04 January 2022*
Published: *16 March 2022*

Citation:
Mandal K, Dutta S, Upadhyay A, Panda A and Tripathy S (2022) Comparative Genome Analysis Across 128 Phytophthora Isolates Reveal Species-Specific Microsatellite Distribution and Localized Evolution of Compartmentalized Genomes. Front. Microbiol. 13:806398. doi: 10.3389/fmicb.2022.806398

Phytophthora sp. are invasive groups of pathogens belonging to class Oomycetes. In order to contain and control them, a deep knowledge of their biology and infection strategy is imperative. With the availability of large-scale sequencing data, it has been possible to look directly into their genetic material and understand the strategies adopted by them for becoming successful pathogens. Here, we have studied the genomes of 128 *Phytophthora* species available publicly with reasonable quality. Our analysis reveals that the simple sequence repeats (SSRs) of all *Phytophthora* sp. follow distinct isolate specific patterns. We further show that TG/CA dinucleotide repeats are far more abundant in *Phytophthora* sp. than other classes of repeats. In case of tri- and tetranucleotide SSRs also, TG/CA-containing motifs always dominate over others. The GC content of the SSRs are stable without much variation across the isolates of *Phytophthora*. Telomeric repeats of *Phytophthora* follow a pattern of $(TTTAGGG)_n$ or $(TTAGGGT)_n$ rather than the canonical (TTAGGG)n. RxLR (arginine-any amino acid-leucine-arginine) motifs containing effectors diverge rapidly in *Phytophthora* and do not show any core common group. The RxLR effectors of some *Phytophthora* isolates have a tendency to form clusters with RxLRs from other species than within the same species. An analysis of the flanking intergenic distance clearly indicates a two-speed genome organization for all the *Phytophthora* isolates. Apart from effectors and the transposons, a large number of other virulence genes such as carbohydrate-active enzymes (CAZymes), transcriptional regulators, signal transduction genes, ATP-binding cassette transporters (ABC), and ubiquitins are also present in the repeat-rich compartments. This indicates a rapid co-evolution of this powerful arsenal for successful pathogenicity.

Whole genome duplication studies indicate that the pattern followed is more specific to a geographic location. To conclude, the large-scale genomic studies of *Phytophthora* have thrown light on their adaptive evolution, which is largely guided by the localized host-mediated selection pressure.

Keywords: *Phytophthora*, genome annotation, effectors, RxLRs, simple sequence repeats, motif preference, whole genome duplication, two-speed genome

INTRODUCTION

Oomycetes are one of the most devastating groups of plant pathogens, resembling mostly filamentous fungi. In the post-genomics era, they are placed under stramenopiles that largely include brown algae and diatoms (Kamoun et al., 2015; Derelle et al., 2016; Hannat et al., 2021). A deep knowledge of oomycetes is very important because of its diversity on host preference involving agricultural crops that causes huge economic loss (Marano et al., 2016). *Phytophthora,* the notorious causal agent for the infamous Irish potato famine, is the most common and pathogenic genus of oomycetes that have more than 180 formal species and are abundant in almost all ecosystems (Yang et al., 2017). They are usually soil-borne in nature and have a wide host range causing the root rot, stem rot, blight, and fruit rot of herbaceous and woody plants (Dodds and Rathjen, 2010).

Bioinformatics tools help in assigning functions to the raw genome sequences (Abril and Castellano Hereza, 2019). For large datasets, assigning important features and meaningful biological information to the sequenced genomes helps to characterize them quickly (Stein, 2001). Due to the complexity of the eukaryotic genomes, gene finding is quite a complicated task compared to prokaryotic genomes (Salzberg, 2019). Moreover, for a non-model organism, it is much more difficult due to lack of trained gene models (da Fonseca et al., 2016). Many annotation pipelines are available recently, and the notable ones such as BRAKER2 (Brůna et al., 2021), funannotate[1], and MAKER (Cantarel et al., 2008) require proper training datasets for predicting gene models. These platforms are mature in predicting gene models and identifying features, but the outcome is extremely unreliable if a trained species dataset is not available. Funannotate, for instance, produces a very poor result if the modeled organism is different from the organism used for prediction. BRAKER2, on the other hand, uses a protein dataset for training and predicts genes using GeneMark-EP + and AUGUSTUS. Easy dissemination of annotation results through data warehouses is also an important area. Over the years, several oomycetes genome resources such as eumicrobedb.org (Panda et al., 2018) and FungiDB (Basenko et al., 2018) have been created and maintained by community members. While both eumicrobedb and FungiDB are primarily based on the genome unified schema (Clark et al., 2005), the portability and ease of handling data are very difficult. Therefore, there is a need for creating resources that can be easily updated.

In *Phytophthora*, genome evolution is influenced by transposable elements (TEs) that give rise to genome fluidity. The genes responsible for pathogenicity, especially the effectors, tend to be localized in TE-rich regions that are gene-sparse regions contrary to the core gene regions, i.e., gene-dense regions (Raffaele and Kamoun, 2012; Engelbrecht et al., 2021). This partitioning of genomes with different evolutionary rates refers to the "two-speed genome" concept.

The mechanisms behind the acquisition and evolution of virulence of *Phytophthora* species is the most studied area for oomycetes biologists. Virulence is controlled by genome architecture, and it subsequently influences the number of specialized effectors that enter the host cells to establish infection (Tyler et al., 2006; Franceschetti et al., 2017). Effectors are broadly categorized into extracellular and intracellular types (McGowan and Fitzpatrick, 2020). Extracellular effectors are secreted on to the host cell and interacted with apoplastic proteins of the host. This includes cell wall-degrading enzymes, protease inhibitors, and elicitins (McGowan and Fitzpatrick, 2017). Intracellular effectors are translated into the host cell and interact with defense-related proteins in the host to manipulate the host immunity. Intracellular effectors are mainly divided into two families, RxLRs (arginine-any amino acid-leucine-arginine) and crinklers (CRN) (Haas et al., 2009). RxLRs are the most abundant family of effectors and are characterized by the presence of conserved amino acid motif arginine (R)–any amino acid (X)– leucine (L)–arginine (R), usually followed by dEER (aspartate glutamate glutamate arginine) domain at their N-terminus (Jiang et al., 2008; Birch et al., 2009). These RxLR-dEER conserved motifs actively participate in translocation, the secretion of effectors into the host cell during the biotrophic phase of infection (Whisson et al., 2007; Schornack et al., 2010; Wawra et al., 2017). The CRN effectors are named after their crinkling and necrosis-influencing activities inside the hosts, which contain conserved LFLAK (leucine phenylalanine leucine alanine lysine) domain at their N-terminal that is associated with translocation into the host cell (Schornack et al., 2010). Effector prediction is a challenging task since they do not generally share features with other protein-coding genes. This is more so since these genes undergo rapid evolution, and therefore, the sequences do not generally possess significant similarity. Another challenge in the prediction of these effectors is due to their location in the genome. Since these effectors are localized mostly at the repeat-rich regions, this region is either unsequenceable or contributes to genome assembly fragmentation. The prediction of effectors from *Phytophthora* is a significant challenge that may lead to the development of disease control strategies.

Simple sequence repeats (SSRs) or microsatellites are the repetition of specific nucleotides distributed in different parts

[1]https://github.com/nextgenusfs/funannotate

of the genomes and are considered as one the most powerful molecular tools for the identification of inter- and intraspecific variability of genomes (Ellegren, 2004; Gonthier et al., 2015; Biasi et al., 2016). The presence of high DNA replication error and mutation rate in the SSR region in comparison with other parts of the genome produce a high degree of length polymorphism within close organisms (Selkoe and Toonen, 2006; Mascheretti et al., 2008; Olango et al., 2015). The distribution of SSRs in the genome is not random and shows preference toward non-coding regions rather than coding regions due to the selection pressure against frameshift mutations. Although exception is applicable for tri- and hexanucleotide motifs because they do not make frameshift, which supports the fact that the distribution of trinucleotide motif SSRs are higher in coding regions (Tóth et al., 2000; Srivastava et al., 2019). There are many important biological features attributed to SSRs such as codominance, multiallelic nature, and genetic markers just to name a few. These features are used for determining the mating type, genome reconstruction, disease dynamics, and determination of population structures (Schena et al., 2008; Biasi et al., 2016; Stewart et al., 2016; Engelbrecht et al., 2017; Parada-Rojas and Quesada-Ocampo, 2018; Hieno et al., 2019; Zhang et al., 2019; Guo et al., 2021). The available literature indicates studies involving fewer species and lacks a broader view.

In the present study, we have selected 128 assembled *Phytophthora* genomes of reasonable quality from Genbank and annotated 70 genomes whose annotation was not available in Genbank using BRAKER2 (Brůna et al., 2021). We have further compared all the species among themselves using several approaches including genome Mash distance, effector clustering, SSR properties, and whole genome duplication. We have compared the outcomes of SSR clustering and effector clustering with the phylogeny that was already described in Yang et al. (2017) and concurred with the finding. To investigate the evolutionary concept of effector localization in *Phytophthora*, we have performed "two-speed genome" analysis. We have gone beyond to check the other genes apart from effectors, which are localized in gene-sparse regions. We have carried out the two-speed genome analysis using the intergenic distances as a function.

We have further created browsable genome resources and repositories of annotated files to bridge the knowledge gap. We have also shared the annotation and training resources for the ease of annotation of the related species. The genome database is available at www.eumicrobedb.org:3000. All the annotated files such as GFF3, coding sequence (CDS), and protein are deposited into https://zenodo.org/record/5785473#.YcB4D2hBzIV and are now made publicly available. All the scripts and programs used in this study are deposited in https://github.com/computational-genomics-lab/scripts-for-SSR-project.

MATERIALS AND METHODS

Data Collection

The FASTA files of all the publicly available *Phytophthora* genomes of reasonable quality were collected from the Genbank file transfer protocol (FTP) site[2]. A total of 162 genomes were available in the National Center for Biotechnology Information (NCBI) as of July 11, 2021. For analysis, we only took genomes having scaffold level assemblies counting to 128 genomes belonging to 33 species (**Supplementary File 1**). The assemblies of several genomes were extremely fragmented, resulting in a difficulty in analysis. Therefore, two genomes, *Phytophthora cambivora* isolate: CBS114087 and *P.* x *alni* were not used for further annotation. This made the total number of annotated species to 31 (yellow marked rows in **Supplementary File 1** are unannotated). We have created standard genome prefixes by using the first three letters of the genus name ("Phy" in this case) and first two letters of species name (e.g., "so" for sojae), followed by the isolate name separated from the genome prefix with an underscore ("_" For example, Phyin_T30-4 stands for *P. infestans* isolate T30-4).

Genome Completeness Prediction

Genome assembly completeness was performed for each of the genomes with benchmarking universal single-copy orthologs analysis (BUSCO v. 5.2.2) (Seppey et al., 2019) using BUSCO data set "stramenopiles_odb10.2019-11-21" containing 100 conserved genes. We have used the genomes as well as predicted proteins as the input for BUSCO analysis. The following command was used to run BUSCO:

```
busco -i <genome input dir> -o <output dir> -m geno -l
stramenopiles_odb10 -c 80 2>&1 | tee <logfile name>
```

Gene Prediction and Genome Annotation

Out of the 128 available genomes of *Phytophthora*, 56 already had annotations available from Genbank resources. We have therefore carried out gene model prediction using the BRAKER2 pipeline (Brůna et al., 2021) for the 72 remaining isolates. Out of the 72 remaining isolates, *P. cambivora* isolate: CBS114087 and *P.* x *alni* could not be annotated due to fragmented genome assembly.

The draft genome assemblies were first cleaned, followed by doing a soft-masking using redmask v 0.0.2[3]. Then, the soft-masked assemblies were used for gene prediction using BRAKER2. For training, we have used protein files for generating training models and, subsequently, genes were predicted. For the species isolates from NCBI, which have well-annotated protein data, we have used one of the isolate protein data as a protein hint file for the gene prediction of other non-annotated isolates of the same species. For example, for the isolates of *P. capsici*, Phyca_CPV-219.fna, Phyca_CPV-262.fna, etc., we have used the annotated proteins of Phyca_LT1534-B.fna as a hint file. On the other hand, for those species that do not have any annotated isolates available, we have merged all the proteins of previously mentioned 56 annotated genomes and used the merged proteome "PROTHINT.faa[4]" as a hint file.

[2]https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/4783/?utm_source=assembly&utm_medium=referral&utm_campaign=KnownItemSensor:taxname

[3]https://github.com/nextgenusfs/redmask

[4]https://zenodo.org/record/5785473#.YcB4D2hBzIV

Downstream Functional annotation was carried out using funannotate[5] that includes several databases such as pfam v34.0, uniprot v2021_03, buscoalveolata_stramenophiles, emapper-2.1.4-2-6-g05f27b0, signalp v5.0b, merops v12.0, and CAZy. The completeness of predicted protein sets was further evaluated on BUSCO v5.2.2 using the stramenopiles_odb10 dataset in "prot" mode.

Phylogenetic Analysis of the Genomes

A phylogenetic tree spanning 128 *Phytophthora* isolates was constructed using Mash distances with the help of Mashtree v.1.2.0 software (Katz et al., 2019). The Mashtree tool used Mash (Ondov et al., 2016) to create MinHash sketches (number of hashed kmers) of the genomes with the help of Mash sketch function with default parameters. Mash distances were then calculated between those sequences using their MinHash sketches, which estimate the mutation rate between them. The more similar genome sequences were likely to share more common MinHashes and less Mash distances. Furthermore, Mash distances were stored in a pairwise distance matrix that was used for building the dendrogram. The neighbor-joining (NJ) algorithm was implemented here. Bootstrapping was performed 1,000 times. The tree was visualized and annotated with the help of iTOL v6 (Letunic and Bork, 2021). The following command was used to run Mashtree.

```
mashtree_bootstrap.pl –reps 1000 –numcpus 50 *.fna – –
min-depth 0 > mashtree.bootstrap.dnd
```

We have now presented the tree in a circular format with certain annotated features. The features include host preference (curated from available literature), predicted genome features such as genome size (in MB X10 i.e., 100 KB range), the number of predicted effectors, and number of SSRs motifs.

Whole Genome Duplication Analysis

Ks (or dS) refers to the expected number of synonymous substitutions per synonymous site, also known as synonymous distances between two DNA CDSs. For the detection of whole genome duplication (WGD) events, the whole paranome Ks distributions, constituting all estimated Ks values for all gene duplication events of the genome, were constructed for each of 126 *Phytophthora* genomes using the CDSs with the help of a whole genome duplication-detecting tool wgd v1.1 (Zwaenepoel and Van de Peer, 2018). We have used the kernel density estimation (KDE) model to the Ks distributions and looked for the peaks in the distributions as an evidence of WGD event.

Genome Binning and Calculation of Flanking Intergenic Regions

We have used in-house perl scripts for determining the intergenic distances flanking the genes in their 5′ and 3′ regions from the BRAKER2-generated GFF3 files or Genbank-annotated files merged with the predicted effector GFF3 files. We have computed the mean, max, and min distances using in-house perl scripts. Most of the effectors are predicted from small fragmented scaffolds; therefore, they had an intergenic distance of zero. We have eliminated such cases from analysis. We have also removed cases where the BRAKER2-predicted gene model overlapped with our predicted effector, resulting in a negative distance between the genes. For computing the significant differences between the 5′ and 3′ Flanking Intergenic Regions (FIRs) of all genes, BUSCO genes, and RxLRs, we have used a two-tailed *t*-test for paired samples. We have plotted the data using Python and R scripts. The scripts are available at: https://github.com/computational-genomics-lab/scripts-for-SSR-project.

Genome-Wide Simple Sequence Repeat Identifications

For SSR identification, we have taken the exact repetition of motif without any mismatch (perfect SSR), which was identified from the whole genome sequence of all downloaded data using the software package GMATA (genome-wide microsatellite analyzing toward application)[6] (Wang and Wang, 2016). We have used 2–10 bp motifs for consideration that are repeated at least five times from both the strands. The following command was used to run GMATA:

```
perl gmat.pl -i *.fna -r 5 -m 2 -x 10
```

Here, -r 5 implies at least five times repetition; -m 2 and -x 10 indicate the minimum and maximum range of SSR motifs. This has generated four different types of output files with different extensions; *.fms file containing formatted sequences used for SSR identification; *.fms.sat1 file containing a statistical summary of the input sequence(s); *.ssr file with tab-delimited text containing the name and length of the scaffold, start–end position of SSRs, number of repetitions and the corresponding SSR motifs; .sat2 file containing the overall statistics of predicted SSRs. Here in this study, we have referred to the .ssr and .sat2 files as SSR and SAT2 files, respectively.

Calculation of Basic Simple Sequence Repeat Features

The basic statistics of SSR analysis derived from the GMATA-generated SSR and SAT2 files were used for comparative studies like SSR frequency, GC content, density, and SSR coverage by using in-house Python scripts available at https://github.com/computational-genomics-lab/scripts-for-SSR-project. Here, SSR density implies the number of bases covered by SSRs per Mb of the genome. SSR coverage denotes the percentage of genome covered by SSRs.

Calculation of In-Frame Frequency of Trinucleotide Simple Sequence Repeat Motifs

In order to identify the SSRs predicted from the whole genome sequences intersecting with the coding regions, we had intersected the SSR files with their corresponding GFF3 files

[5]https://github.com/nextgenusfs/funannotate

[6]https://sourceforge.net/projects/gmata/

using bedtools v2.26.0 (Quinlan and Hall, 2010). From the bed-intersected files, we have collected the gene IDs that overlapped with the SSR regions. Then, we have calculated the in-frame frequency of the SSR trinucleotide motifs present within the CDS of the same gene. Thereafter, we have calculated the cumulative frequency for each trinucleotide SSR motif across the particular genome using an in-house Python script "SSR_CDS_overlap.py" available at https://github.com/computational-genomics-lab/scripts-for-SSR-project. Finally, we have generated a heatmap based on these values. We have also computed the abundance of trinucleotide SSR motifs exclusively within the coding regions using the GMATA software with the predicted CDSs of 126 annotated genomes as input files. A heatmap was generated using these values subsequently to show the abundance of specific trinucleotides within the coding regions.

Effector Prediction

A basic pipeline (**Supplementary Figure 1**) was created for the effector prediction of *Phytophthora*. At first, all the possible open reading frames (ORFs) within the length 150–1,500 nucleotides were extracted from all the six frames of the assembly files using the getorf tool of EMBOSS package v. 6.6.0.0 (Rice et al., 2000). The extracted ORFs were translated in one frame using the transeq tool of the EMBOSS package. These translated sequences were first classified as secretory proteins based on their SignalP v. 5.0b (Armenteros et al., 2019) scores in the N terminus. The secretory proteins were screened for the presence of any transmembrane helices (TMHs) using TMHMM Server v. 2.0 (Krogh et al., 2001). The SignalP containing proteins lacking any TMH was passed through TargetP-2.0 analysis (Armenteros et al., 2019). TargetP predicts the presence of N-terminal pre-sequence based on where they are targeted including signal peptide (SP) (responsible for secretion), mitochondrial transit peptide (mTP), chloroplast transit peptide (cTP), or thylakoid luminal transit peptide (luTP). The ones containing mTP, cTP, and luTP are removed at this step.

EffectorO (Nur et al., 2021) was used to identify the putative effector proteins from the secretome with default parameters. RxLR hmm model "pf16810.hmm" was downloaded from the pfam database (06/08/2021)[7]. An hmmsearch was performed against the pf16810.hmm database for the effectors predicted by the EffectorO program to ascertain the presence of RxLR motifs. For CRN effector prediction, *Phytophthora*-specific CNR proteins were retrieved from NCBI database, followed by multiple sequence alignment using MUSCLE v3.8.31 (Edgar, 2004). A CRN-specific hmm model was built using the hmmbuild tool from HMMER 3.1b2 package[8]. To detect the effectors containing the WYL (tryptophan tyrosine leucine) domain, "WYL_3.hmm" (downloaded on 06/08/2021) from https://pfam.xfam.org/family/PF18488 database was used.

Orthology Analysis of RxLRs

The outputs from the effector prediction analysis pipeline (**Supplementary Figure 1**) resulted in a total number of 19,269 RxLR effectors across 128 isolates. Proteinortho v5 (Lechner et al., 2011) tool was used with default parameters for identifying the clusters (95% minimum reciprocal similarity for additional hits; the E-value of 1e-05 for the blastp; minimum percent identity of 25 for best blast hits; minimum coverage of 50% for the best blast alignments). An unweighted pair group method with arithmetic mean (UPGMA)-based species tree was generated for RxLR orthologous clusters of 128 isolates with the help of po2tree.pl program, which is provided with the Proteinortho v5 tool. The tree was further visualized using iTOL v6[9] (Letunic and Bork, 2021).

Browsable Annotated Component Development

At Indian Institute of Chemical Biology (IICB), we have created a React-based single-page application. The React framework was chosen as it is easy to create independent reusable components. New functionalities and plugins are easier to incorporate with this framework. The app, which is in its testing version, includes a (Buels et al., 2016) plugin. The app is available at this address: www.eumicrobedb.org:3000.

Statistical Analysis for Correlation

To establish the correlation between various numerical variables, scatterplots were made using R programming language (Ihaka and Gentleman, 1996). The packages used were "ggplot2" (Wickham, 2011) to plot the scatterplot and "ggpubr" (Kassambara and Kassambara, 2020) to add the Pearson correlation coefficient as well as the *p*-value of the scatterplot. The script for doing so is present in the following Github link: https://github.com/computational-genomics-lab/scripts-for-SSR-project under the folder R-scripts.

RESULTS AND DISCUSSION

Number of Predicted Genes Are Co-related With the Genome Size

Out of the 128 genomes studied, 126 genomes had more than 90% BUSCO completeness. However, two species, e.g., *P. cambivora* isolate: CBS 114087 and *P.* x *alni* had extremely fragmented assemblies with 72,332 and 1,184,74 scaffolds, respectively. This exceeded the number of sequences that can be handled in BRAKER2; therefore, we could not annotate these two strains. The genome size and the number of predicted genes in the *Phytophthora* species are correlated ($r = 0.409$, p-value 0.00001). The highest number of proteins were predicted (36,721) from *P. syringae* BL57G having a genome size of 74.93 Mb, whereas the lowest was predicted from *P. kernoviae Chile4*, Phyke_Chile4 with a genome size of ∼37 Mb. Other isolates of *P. kernoviae* remain the genomes with the least number of genes (<10,500) that also had the least genome size (<37 Mb).

Since the genomes studied were near 90% complete, it is fair to assume that the number of predicted genes represent

[7]http://pfam.xfam.org/family/PF16810

[8]http://hmmer.org/

[9]https://itol.embl.de/

90% of the genes. *P. kernoviae* have the least genome sizes (36–38 Mb) followed by *P. ramorum* species. Both *P. kernoviae* and *P. ramorum* infect tree species and are mostly homothallic and biotrophic[10, 11]. The species with larger genomes are *P. infestans* and *P. cambivora*. Both the species are heterothallic and have undergone transposon mediated genome expansion (Haas et al., 2009). While it is difficult to establish the link between the genome size and the virulence, it is a well-established fact that organisms having larger genomes and larger gene repertoire, is due to their heterothallic nature. While biotrophs such as *Hyaloperonospora arabidopsidis* are known to have streamlined genomes (Baxter et al., 2010), hemibiotrophs have larger genomes.

Genome Size and Number of Simple Sequence Repeats Are Positively Correlated

We have carried out SSR finding with 2–10 bp units repeated at least five times in all the 128 genomes. This resulted in 391,318 SSRs. The number of SSRs in the genomes is positively correlated with the genome size (**Supplementary Figure 2A**, $R = 0.84$, $p = 2.2$ e-16). All the SSR files containing microsatellite data and SAT2 files containing information regarding SSR statistics are publicly available at https://zenodo.org/record/5785473#.YcB4D2hBzIV. Studies with mosquitos and other species also reveal that SSR frequencies are directly co-related with the genome size (Srivastava et al., 2019). In order to rationalize the number of bases in SSRs per Mb of genome, we have computed the SSR density. *P. agathidicida* isolate-NZFS3770 had the lowest SSR density (921.81 bp/Mb of genome) and *P. boehmeriae* isolate SCRP23 had the highest density (152,348.35 bp/Mb of genome). The genome sizes of *P. agathidicida* NZFS3770 and *P. boehmeriae* SCRP23 are 37.23 and 39.96 Mb, respectively, which is low compared to other *Phytophthora* genome sizes (**Supplementary File 1**). Contrary to the number of SSRs and genome size correlation, the genome size and density of SSRs have very little correlation (**Supplementary Figure 2B**, $R = -0.19$, $p = 0.034$). The percentage of SSRs per genome or genome coverage is another way of depicting SSR density. As expected, *P. boehmeriae* isolate SCRP23 has 15.23% of the genome covered with SSRs. This is followed by *P. ramorum EU* isolates (11%–12%). *P. agathidicida* NZFS3770 has lowest coverage with 0.09% (**Figure 1A** and **Supplementary File 2**). Among the *P. ramorum* isolates, the ones isolated from EU (European Union) had significantly higher number of SSRs than that of the American isolates, NA1 strain Pr102 and CDFA1418886 (11%–12% vs. 7%). The European strains of *P. ramorum* are more aggressive than the original NA1 strains found infecting coastal districts of California, United States. Since microsatellites mutate 10 orders of magnitude greater than commonly occurring point mutations (Gemayel et al., 2012), it could be speculated whether increased frequencies of SSRs in virulent isolates indicate greater adaptability. Gain and loss of gene functions are attributed due to frameshift mutations and subsequent fixation. The presence of a higher number of SSRs could possibly mean that the genomic region is in a state of flux and may contribute to adaptation, leading to increased virulence. On the contrary, *P. agathidicida* isolate NZFS3770 having the least SSR density is an extremely virulent pathogen in Kauri (*Agathis australis*) (Studholme et al., 2016).

[10]https://www.cabi.org/isc/datasheet/40972

[11]https://www.cabi.org/isc/datasheet/40991

The Lower-Order Simple Sequence Repeat Motifs Represent the Major Class in All Isolates

The dinucleotide SSRs are the most abundant class with an average of 66.10% followed by trinucleotide (29.42%) and tetranucleotide SSRs (2.35%). Altogether, the di- tri-, and tetra nucleotide SSRs constitute 97.89% of total SSRs, while the remaining are the penta- to decanucleotide repeats. For all the species, the abundance rank of SSRs are always dimer > trimer > tetramer. The higher-order SSRs do not follow any trend and are mostly species and isolate dependent (**Supplementary File 3**). The average number of SSR motifs across the isolates is 154.81 with the lowest and highest number of motifs in *P. palmivora* isolates B4_PPRK (90) and *P. cactorum* isolate P404 (286), respectively. It was noticed that the number of SSR motifs are negatively correlated with their genome size, although the correlation is not strong (**Supplementary Figure 2C**, $R = -0.12$, $p = 0.18$). Out of 128 genomes, 54 had no decanucleotide SSRs, 37 had no octameric SSRs, and 3 without heptameric SSRs.

SSR motif lengths were classified into two categories, Class I (SSR length $\geq$ 20 base pair) and Class II (SSR length < 20 base pair) and presence of Class I SSRs indicate hyperpolymorphism. We found that all genomes currently studied show more than 80% class II category SSR. Class II SSRs tend to be less variable compared with Class I due to low probability of slipped-strand mispairing on the short SSR strand (Temnykh et al., 2001). Thus, Class I SSRs are better for polymorphism identification than Class II. So, designing markers from Class I SSRs for the identification of polymorphism among the genotypes of a species could be more reliable.

In order to establish the relationship between genomic GC content and SSR GC content, we have computed the GC content for genomic DNA and SSR regions of the genome. While the genomic region had 52.13 ± 1.18% GC, SSR region had comparatively lower 48.22 ± 1.5% GC content (**Figure 1B**). Statistical analysis shows that there was positive correlation between genomic and SSR GC content (**Supplementary Figure 2D**, $R = 0.39$, $p = 7.1$ e-06). The GC contents of the *Phytophthora* SSRs are in fact a characteristic feature for the species. Chlorophytes are characterized by GC-rich SSRs; most fungal species are reported to have intermediate GC- containing SSRs, while complex genomes such as plants carry high AT content in SSR (Srivastava et al., 2019).

In order to establish the relatedness of species and clade (Yang et al., 2017) on the basis of SSRs, we have performed principal component analysis (PCA) for di-, tri-, and tetranucleotide motifs from all the genomes. Results indicate a strong species-specific distribution of SSR motifs. PCA shows three clusters (**Figure 1C**). Clade 1 and clade 4 are present within cluster-A, which indicates that they are close to each other. This fact was already established

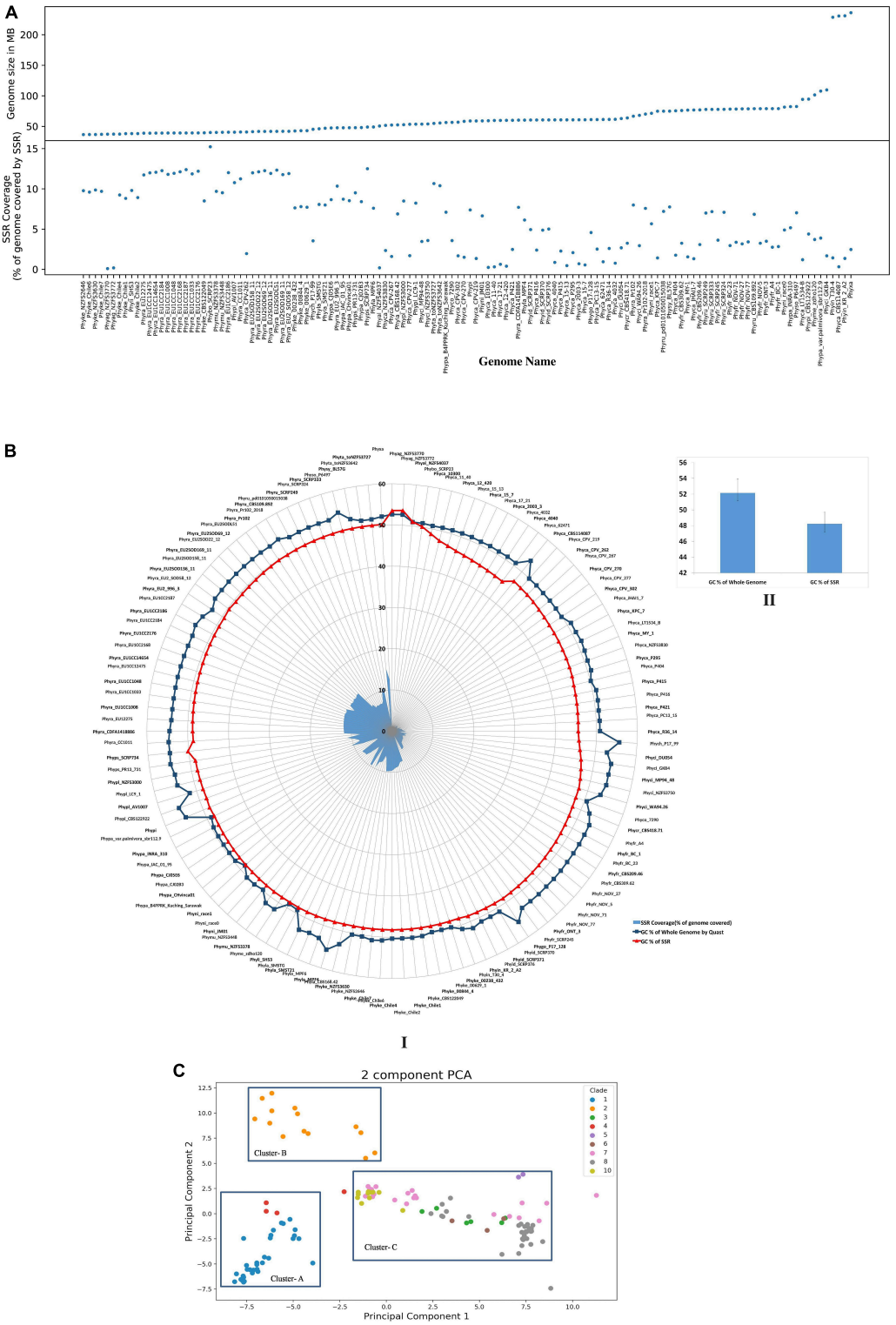


FIGURE 1 | Comparison of various SSR attributes within genomes. **(A)** Comparison of genome size and SSR coverage. **(B)** GC content of SSRs and their corresponding genomes. Genomic GC content of *Phytophthora* (52.13 ± 1.18%) and GC content of SSRs is (48.22 ± 1.5%). I: Is the circular representation of GC content. II: Represents the bar plots of mean genome GC content and SSR GC content. **(C)** PCA clustering using SSR motifs. Each dot represents a genome and is colored based on their clades as described in Yang et al. (2017). Di, tri, and tetramer motifs were taken from each genome, and PCA was done. Genomes from each clade clusters together more than others indicate phylogeny-based SSR variation. Cluster-A contains genomes from clade 1 and 4. Clade 2 positioned separately as Cluster-B. The rest of the genomes from clades 3, 5, 6, 7, 8, and 10 are clustered together as Cluster-C.

from phylogenetic analysis, which was based on seven nuclear genetic markers as described by Yang et al. (2017). Only clade 2 represents cluster-B. Cluster-C contains the highest number of isolates and represents clades 3, 5, 6, 7, 8, and 10. From the clustering, it was demonstrated that clade 2 has more distinct SSR motifs than others. Our analysis indicates that the SSR composition of clades 3, 5, 6, 7, 8, and 10 (Cluster C) is closer with each other than with other clades (Cluster A and Cluster B).

"TG/CA" Dinucleotide Motifs Represent the Most Abundant Class of Simple Sequence Repeats Across All *Phytophthora* Genome Isolates

The occurrence of dinucleotide, trinucleotide, and tetranucleotide SSRs were plotted using heatmap, which constitutes 97.89% of entire predicted SSRs' cumulative length (**Figure 2A**). For this, we took two complementary motifs as groups since strandedness is unknown. We have calculated the percentage of the motifs for each class e.g., di-, tri- and tetra-. For example, *P. agathidicida* isolate NZFS3770 contains 1,170 dimeric SSRs and the TG/CA motif is present 294 times, therefore making it 25.12% of the total number of dimers. The dinucleotide SSR motifs are often used as molecular markers due to their higher mutation rates than other types of SSRs (Karaoglu et al., 2005). TG/CA is the most preferred di nucleotide motifs (more than 23% in average among all dinucleotides) as well as among all SSR motifs (more than 15% in average). This pattern was observed in all the 33 species without a single exception; so, it can be concluded that TG/CA motifs are the characteristic features of genus *Phytophthora*. In several studies, species-specific preference for a particular motif for a genus like *Fusarium, Aspergillus,* and *Nicotiana* has been discussed (Mahfooz et al., 2015, 2016; Wang et al., 2018). The second highest percentage of motifs are AC/GT and AG/CT, which occupied 17.86 and 17.24% respectively. This is followed by GA/TC (15.79%), AT/AT (7.55%), GC/GC (7.54%), CG/CG (6.63%), and TA/TA (4.38%) motifs. Presence of TG/CA and AG/CT in higher percentage may be due to their amenability to low mutation rate (Guo et al., 2009). It has been reported that TG/CA and AC/GT were predominant motifs in the mammalian system and AT/AT and TA/TA were abundant in the plant systems (Lagercrantz et al., 1993; Morgante et al., 2002). Dinucleotide motif composition of *Phytophthora* follows opposite to the plant system and contains low amounts of AT or TA motifs. This attribute has been the basis behind separating the contaminating plant DNA from oomycetes DNA (Tripathy et al., 2012).

Tri and Tetranucleotide Motifs Containing "TG/CA" Patterns Represent the Most Frequent Class of Simple Sequence Repeats

Out of 30 possible groups of "Tri" motifs, 6 motifs (AAG/CTT, AGA/TCT, AGC/GCT, CAG/CTG, GAA/TTC, and GCA/TGC) are present predominantly and occupy 47.1% of total trinucleotide-containing SSRs (**Figure 2B**). CAG/CTG motifs are found to be the dominant class in 20 *Phytophthora* species that represent 11.27% of total trinucleotide SSR motifs. *P. colocasiae, P. idaei, P. multivora, P. palmivora, P. plurivora*, and *P. parasitica* on the other hand have AAG/CTT as the most predominant motif, similar to fungi *Trichoderma atroviride, T. virens, Aspergillus nidulans* and *A. oryzae* (Mahfooz et al., 2016, 2017). *P. capsici, P. megakarya*, and *P. litchii* show GAA/TTC dominance. At the same time, ATG/CAT and ATC/GAT are the lowest common motifs with less than 1% occurrence.

It has been found that occurrence of trinucleotide SSRs on ORF and 5′- UTR regions was much higher than the non-coding regions of the genome (Gonthier et al., 2015). Thus, motif dominance, which represents a complete codon for trinucleotide SSRs, makes sense that it has an important role in molecular mechanism. Amino acids encoded by the most abundant motifs (CAG/CTG) are leucine and glutamine, AAG/CTT encodes leucine and lysine, respectively. In order to find out the presence of trinucleotide motifs in coding region, we have used CDS file as an input for SSR identification and the result exhibited same pattern with dominance of above mentions groups motifs (**Supplementary Figure 3**). The presence of trinucleotide on the coding region is not enough for translation, so, we performed overlap of predicted SSRs falling in-frame with the coding regions. The heatmap of in-frame analysis also shows dominance of CAG, CTG, and AAG motifs (**Supplementary Figure 4**). CAG and AAG codes for leucine amino acid. It has been reported that amino acid leucine helps in zoospore germination, which eventually helps in the establishment of infection to the host of *Phytophthora* (Jiang et al., 2019). Thus, SSRs with leucine CDS may have a vital role in germination, but further in-depth study is required for establishing this link. Basic amino acid lysine and arginine induced encystment in *P. cinnamomi* (Byrt et al., 1982). This may be the reason why these motifs are conserved across all the *Phytophthora* species. Other amino acids encoded by the highly abundant motifs are serine, alanine, arginine, glutamic acid, phenylalanine, and cysteine.

Among the tetranucleotide motifs, GACA/TGTC motif was the most commonly occurring motif among all the 128 genomes studied and occupied 7.45% of the total tetranucleotide SSRs. This was followed by ACAG/CTGT (5.15%), AGTG/CACT (4.16%), and AGAC/GTCT (4.07%) (**Figure 2C**). These observations further strengthen the predominance of TG/CA dinucleotide that forms a part of the tri- and tetranucleotide motifs, representing the major class.

Higher-Order Simple Sequence Repeat Motifs Are Specific to Individual *Phytophthora* Species

The higher-order motifs such as the tetra- to decanucleotide repeats are unique for each of the isolates (**Table 1**). It is also noteworthy that 8 dinucleotide and 22 out of 30 trinucleotide motifs are common in *Phytophthora* genomes and possibly are characteristic features of *Phytophthora* species.

Out of the 128 genomes, 110 genomes have their unique motif containing SSRs, which is specific only to them regardless of

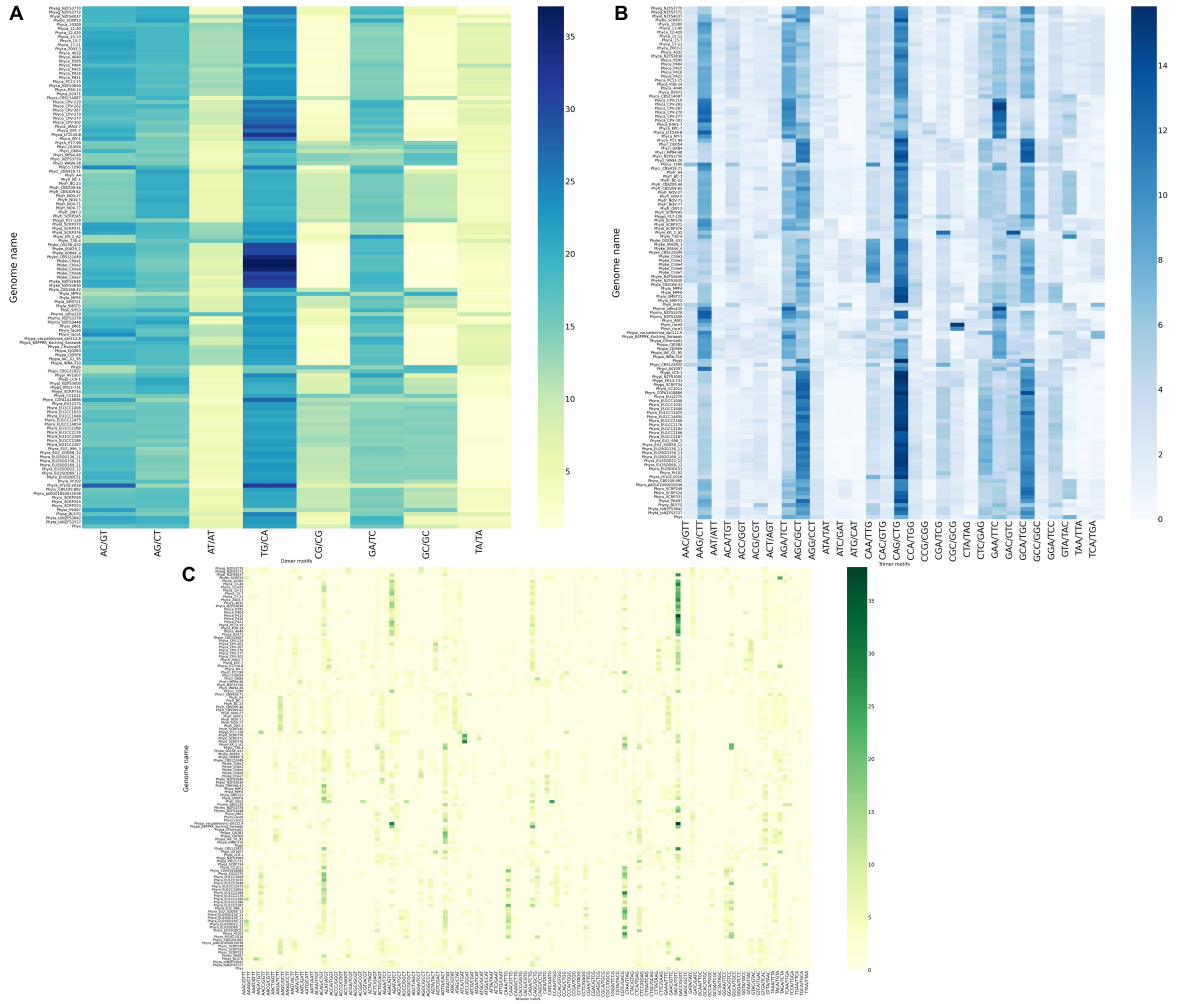


FIGURE 2 | Construction of heatmap of SSR motifs. **(A)** Heatmap of dinucleotide motifs. Out of eight group motifs, TG/CA motif was the most preferred motif in all the genomes and *P. kernoviae* contains the highest percentage of motifs among all the *Phytophthora* genus. Scarcity of AT- containing motif is clearly visible. **(B)** Heatmap of trinucleotide motifs. CAG/CTG motif preference is found in maximum number of genomes. Only a few trinucleotide motifs are predominant, while clearly, the AT-containing motifs such as ATG/CAT and ATC/GAT are very scarce. **(C)** Heatmap of tetranucleotide motifs. There is a clear dominance of GACA/TGTC-containing motif, which has a TG/CA pattern embedded in it.

the species (**Supplementary File 4**). *P. colocasiae* isolate-7290 contains the highest number of unique motifs (81), followed by *P. cactorum* isolate-P404 (46), followed by *P. cinnamomi* isolate-GKB4 (32).

TABLE 1 | Common SSR motifs found across all the *Phytophthora* isolates.

TG/CA	GCT/AGC	CAC/GTG
AC/GT	TGC/GCA	TGA/TCA
GA/TC	TCT/AGA	CCA/TGG
CT/AG	TTC/GAA	TCC/GGA
AT/AT	CTT/AAG	ACC/GGT
CG/CG	CGA/TCG	ACG/CGT
GC/GC	GAG/CTC	GTA/TAC
TA/TA	GTC/GAC	AGG/CCT
CTG/CAG	AAC/GTT	AGT/ACT
CAA/TTG	TGT/ACA	AAT/ATT

For species studied with higher numbers of isolates, the number of species-specific SSR motifs were less and isolate specific SSR motifs were higher (**Supplementary File 5**). For example, in the case of *P. capsici* (10 isolates), *P. fragariae* (11 isolates), and *P. ramorum* (23 isolates) had no common species specific motifs. In case of *P. cactorum* (18 isolates) has only a single common motif. This possibly indicates that SSR markers are more isolate specific. So, a unique SSR for each isolate can be used to design isolate-specific SSR markers.

Telomeres Contain an Extra "T" in Addition to the Canonical Telomeric Repeat TTAGGG

For telomere-like sequence identification, we have manually searched for an ancestral telomere motif TTAGGG by Notepad + + software on the GMATA derived SSR data

with a minimum of three repeats. Interestingly, it was found that the TTAGGG motif was present with an extra T (thymine) on the start or end position of the motif, e.g., $(TTTAGGG)_n$ or $(TTAGGGT)_n$. Most of the time these motifs were located at the start or end regions of a scaffold, indicating the end of the chromosome (**Supplementary File 6**, sheet 1 and sheet 2). A previous study by Fulnečková et al. (2013), reported that $(TTTAGGG)_n$ is a characteristic feature of telomere sequence for plants and oomycetes while mammals and fungi have $(TTAGGG)_n$.

Absence of Core RxLR Clusters Is an Indication of Their Rapid Divergence

The standard effector prediction pipeline (Chepsergon et al., 2021) was used to predict the effectors. The number of proteins retained in each step is shown in **Supplementary Figure 6**. The percentage of RxLR effectors were much higher than other predicted class of effectors as it is primarily associated with infection of host and a recent study also gives same indication (Gao et al., 2021). A significantly higher amount of RxLR effectors containing species are *P. megakarya, P. palmivora, P. cambivora, P. ramorum* (isolate- Pr102 and Pr102-2018), *P. infestans, P. nicotianae, P. sojae*, etc. On the other hand, the lowest number of RxLR was found in *P. pisi* followed by *P. chlamydospora*, *P. syringae* (**Supplementary File 7**). For CRN motifs containing effectors, *P. infestans* and *P. sojae* have the highest number. A higher number of CRN effectors are an indication of preference toward necrotrophic life cycle (Stam et al., 2013).

For the prediction of core RxLR effectors (CRE), we took all the RxLR effectors and ran ProteinorthoV5 through them in order to detect orthologous genes within different species. This resulted in 1,461 orthologous clusters. The singletons containing single RxLRs were discarded and were not considered for further analysis (**Supplementary File 8**). The largest cluster contained 150 proteins representing 105 isolates, which indicates the presence of co-orthologs. We could not predict any core ortholog common across all the 128 isolates studied. Further, we have built a UPGMA species tree on the basis of the clustering patterns of the effectors (**Figure 3**). In most of the cases, the effectors are species specific and isolates of the same species clustered together (**Figure 3**). There are exceptions in cases of *P. ramorum*, where the RxLRs make two distinct groups. Group 1 contains isolates-CDFA1418886, EU1CC1008, Pr102, and Pr102-2018 that are more close to *P. lateralis*, whereas other 19 isolates of *P. ramorum* (Group-2) had more similarity with *P. taxon totara* and *P. syringae*. Another exception is *P. kernoviae* where isolate-Chile 6 and Chile 7 makes a different group from the other isolates of the same species. *P. parasitica* isolate INRA-310 and *P. nicotianae* isolate JM01 make a closer group than other isolates of their own species. The genome Mash distances (**Figure 4**) have grouped the isolates from individual *Phytophthora* species together. However, clusters based on RxLRs grouped isolates of *P. ramorum* into two distinct groups. One containing NA1 isolates from the United States and the other one containing the EU isolates. Similarly, in the case of *P. kernoviae*, two distinct groups were formed. The evolution of the pathogenicity of *Phytophthora* is very complex and driven by many factors that are heavily dependent on host preference. It is therefore not clear if the RxLRs bearing close similarity among other species is an evolutionary strategy for survival.

Flanking Intergenic Region Distance Indicates Clear Two Speed Genome Architecture in All the *Phytophthora* Isolates

We have computed the intergenic distances of the genes in each of the species having predicted gene models (No # 126) (**Supplementary File 9**). The distances and their mean values are provided in **Supplementary File 9**. We have conducted a two-tailed *t*-test for paired samples for comparing the average values of flanking 5′ distance of the genes and 3′distance of the genes. The average 5′ FIRs and the 3′ FIRs among all the species did not have any significant difference (p-value = 0.919; for BUSCO genes = 0.931; for RxLR effectors = 0.81). However, between the 5′ genomic distance of all genes with RxLRs, the p-value is 1.232^{7} and the 3′ distance is 7.93^{-8}. The average 5′ intergenic distance between BUSCO genes and the RxLR have a p-value of 1.68^{-16} and the 3′ distance is 3.68^{-18} (**Supplementary File 9**, **Supplementary Figure 6**, and **Figure 5**). This confirms the two-speed genome theory involving the RxLR effectors.

We have further curated the annotations of the 1,000 genes that are placed at the extremely sparse genomic locations in each of the *Phytophthora* genomes (total # 126 × 1,000 = 126,000 genes). It is interesting to note that out of the 126,000 genes studied, 123,520 had annotations. 42.61% (52,643) are annotated as hypothetical proteins, without any known functions. Among the annotated proteins, 1,004 are RxLR proteins, 255 are CRN, and 263 are elicitins. Among others, the most notable ones are carbohydrate-active enZYmes (CAZymes) such as peptidases (288), pectin esterases (205), and glycosyl transferases (137) (see text footnote 4). Among the other categories are the transcription factors, CW-ype zinc finger protein, CXXC motif-containing genes, EF–hand proteins, PWWP domain- containing protein, calcineurins, ubiquitins, etc. Signal transduction proteins such as WD40 are in high numbers in the gene-sparse regions. We have located thousands of transposons and retroposons in the gene- sparse regions of the 126 *Phytophthora* species.

Numerous reports are available to suggest that the oomycetes pathogens have rapidly evolving powerful arsenals that are used to combat the host defense mechanisms. Pathogens combat hosts with the choicest effectors that are not randomly distributed in the genome. Rather, they are localized in regions rich in transposons and repeats (Dong et al., 2015). Many oomycete organisms have already been studied with a robust two-speed genome composition, where the effectors are located in gene-sparse regions (Vetukuri et al., 2018; Malar et al., 2019). However, there is a lack of extensive studies on the overall composition of the gene-sparse regions. We have analyzed the 1,000 genes of each genome located in the most gene-sparse regions in all the 126 genomes under study. As expected, the genes involving pathogenesis are enriched in this region. Apart from that, many

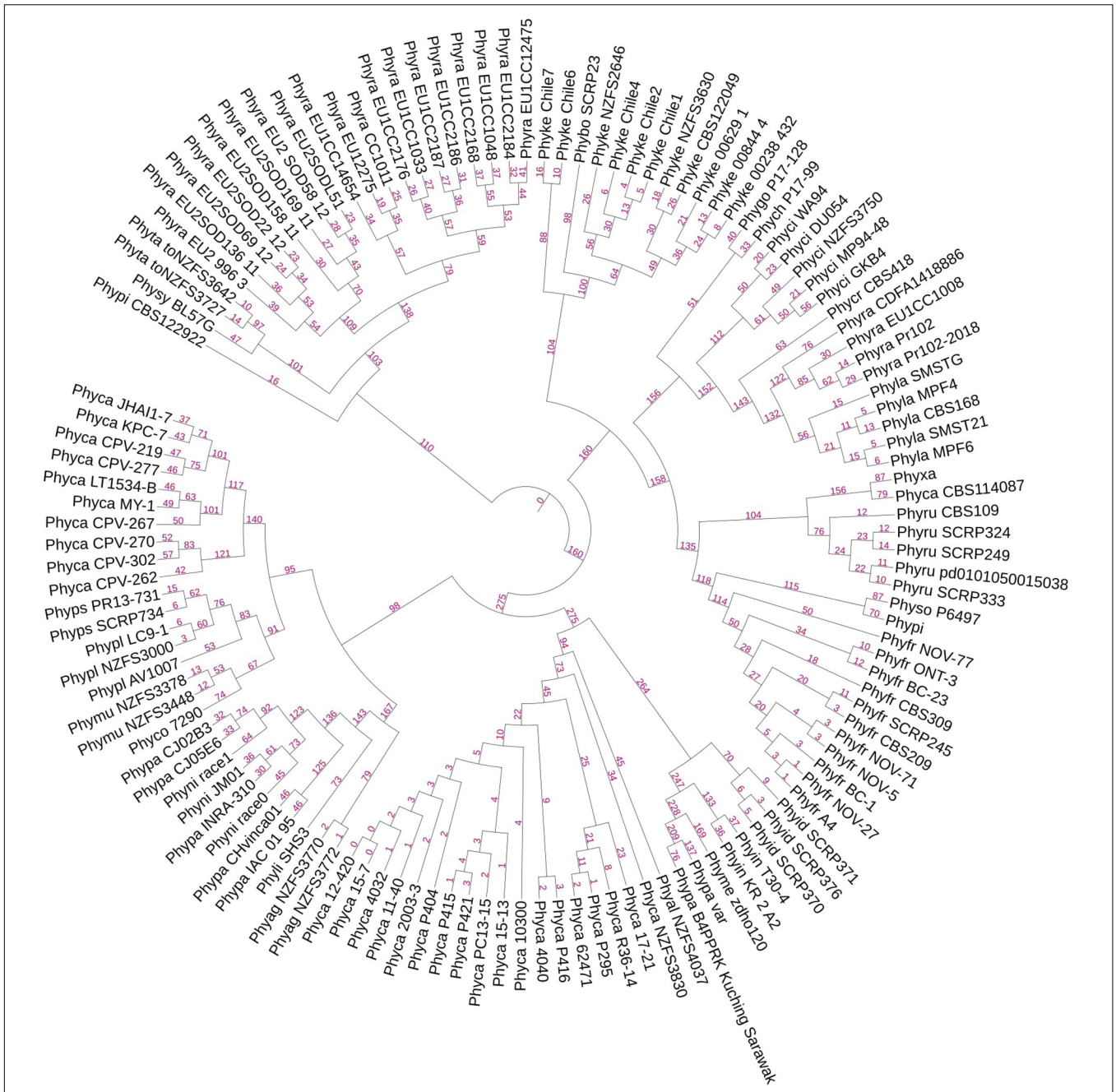


FIGURE 3 | A UPGMA-based rooted species tree showing placement of 128 genomes on the basis of RxLR orthologous clusters. The branching lengths are shown in pink color on the branches of the tree. Species-specific distribution of RxLR effectors was found in most of the cases. There are exceptions where some species have closeness to other species than within themselves. Examples are species *P. parasitica*, *P. nicotianae*, and *P. kernoviae*.

regulatory genes, transcription factors, and signal transduction genes are located in these areas (see text footnote 4).

Two Types of Genome Duplication Events Occur in *Phytophthora*

Genome duplication analysis was conducted using the methods described in Zwaenepoel and Van de Peer (2018). Small-scale duplication and the loss of duplicated copies are not under selection pressure and are a continuous process. However, if large-scale duplication events occur, i.e., whole genome duplication (WGD), then it is visible as peaks in the number of retained duplicates (Zwaenepoel and Van de Peer, 2018). We have used kernel density to show genome duplication events in *Phytophthora* species. Our results show various levels of genome duplication in all the isolates. Interestingly, we have observed two types of genome duplication patterns and we have categorized

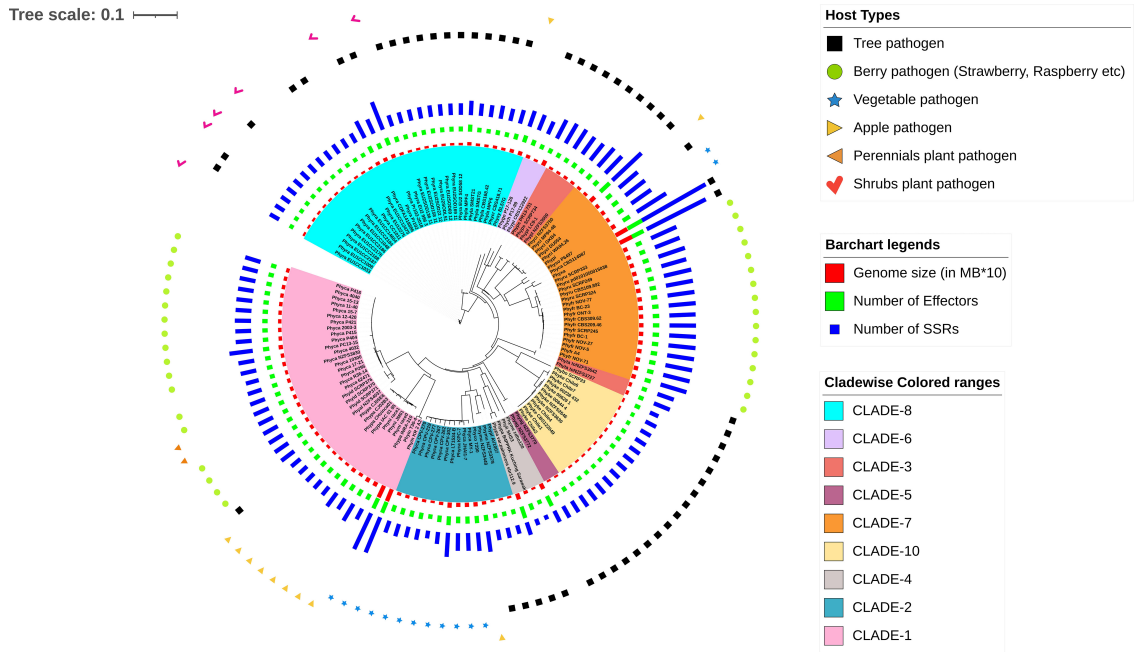


FIGURE 4 | A rooted phylogenetic tree showing evolutionary relationships and the connections between the 128 *Phytophthora* isolates. The tree is generated using NJ methods in the Mashtree software. The genomes are highlighted in different colors according to their clades as described by Yang et al. (2017) and are as follows: cyan, CLADE-8; light violet, CLADE-6; deep pink, CLADE-3; deep violet, CLADE-5; orange, CLADE-7; pale yellow, CLADE-10; gray, CLADE-4; blue, CLADE-2; and light pink, CLADE-1. Different genome features are represented in the form of bar charts such as red bar plot for genome size (in MB*10), green for number of effectors, and violet for number of SSRs. The host types for the *Phytophthora* pathogens are shown using different symbols over the bar charts such as black square for Tree pathogen, green circle for Berry pathogen (strawberry, raspberry, etc.); blue star for Vegetable pathogen; pale-yellow right-sided triangle for Apple pathogen; orange left-sided triangle for Perennials plant pathogen; and pink tick for Shrubs plant pathogen.

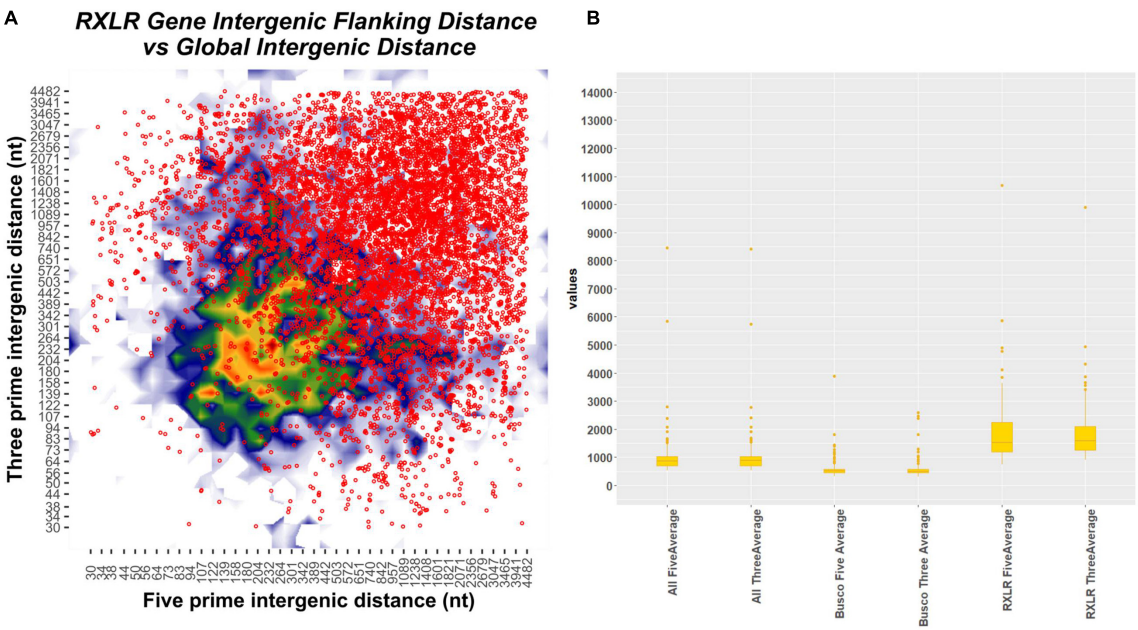


FIGURE 5 | Plot showing flanking intergenic distances of different categories of genes of all the 128 *Phytophthora* species concatenated together. **(A)** Contour plot depicting the five prime and three prime intergenic distances plotted in x- and y-axis. The scatter plot on the top of the contour plot in red dots depicts the position of the RxLR effectors overlaid on the contour plot. There is a clear indication that the RxLRs have significantly large intergenic distances than the other genes. **(B)** Box plot of intergenic distances (both 5′ and 3′ regions) of all the genes concatenated together with BUSCO genes and RxLR genes. The position and the median point of the boxes are indicators of higher 5′and 3′ distances seen in all the *Phytophthora* genomes.

them as type I and type II. In the case of type I, whole genome duplication happens once at K_s 0–0.5, and then the numbers of duplicated genes decrease gradually due to lesser selection pressure, giving an L-shaped distribution to the graph, although a sufficient number of duplicated genes are present (**Figure 6A**). In type II, the pattern of duplication is quite different from type 1 where we have observed the presence of distinct peaks at higher K_s 2.0–2.5. This kind of pattern occurs due to the increase in duplication frequency (**Figure 6B**).

All isolates of *P. capsici, P. cactorum, P. fragariae, P. idaei, P. infestans, P. megakarya, P. palmivora, P. parasitica, P. nicotianae, P. rubi, P. sojae, P. pinifolia, P. colocasiae, P. cryptogea*, and *P. cryptogea* show type I WGD. Previous studies provide evidence for the presence of ancestral WGD in *P. capsici* (Cui et al., 2019), *P. cactorum* (Yang et al., 2018), *P. infestans*, and *P. sojae* (Martens and Van de Peer, 2010). Recently, Morales-Cruz et al. (2020) investigated the whole genome duplication of two species, e.g., *P. megakarya*, and *P. palmivora*. They have demonstrated that both the species go through independent WGDs, which results in large genome size with a higher number of RxLR, CRN, and other pathogenesis-related genes. Type II WGD was noticed in all isolates of *P. agathidicida, P. kernoviae, P. lateralis, P. litchi, P. multivora, P. pisi, P. plurivora, P. pluvialis, P. pseudosyringae, P. chlamydospora*, and *P. taxon totara.* It was also noticed that different isolates of the same species exhibited different types of WGDs. We studied WGD in all 22 isolates of *P. ramorum* where three isolates show type I WGD and rest show type II WGD. The type I WGD-containing isolates of *P. ramorum* are from the United States (Phyra_CDFA1418886, Phyra_Pr102, and Phyra_Pr102-2018), whereas type II-containing isolates are from Europe. We have also shown that the type I-containing ramorum isolates have a lesser number of SSRs than the type II-containing

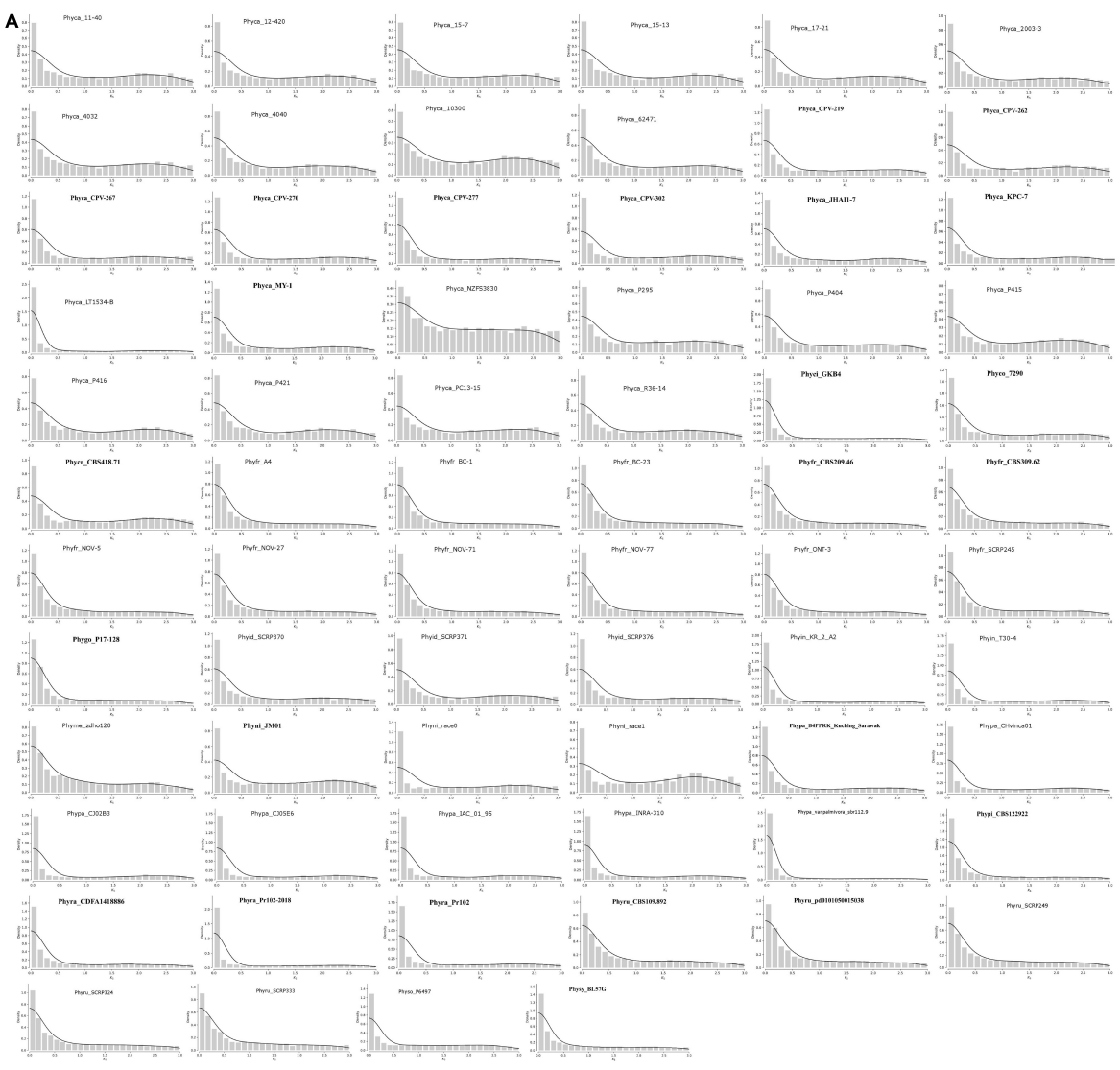


FIGURE 6 | (Continued)

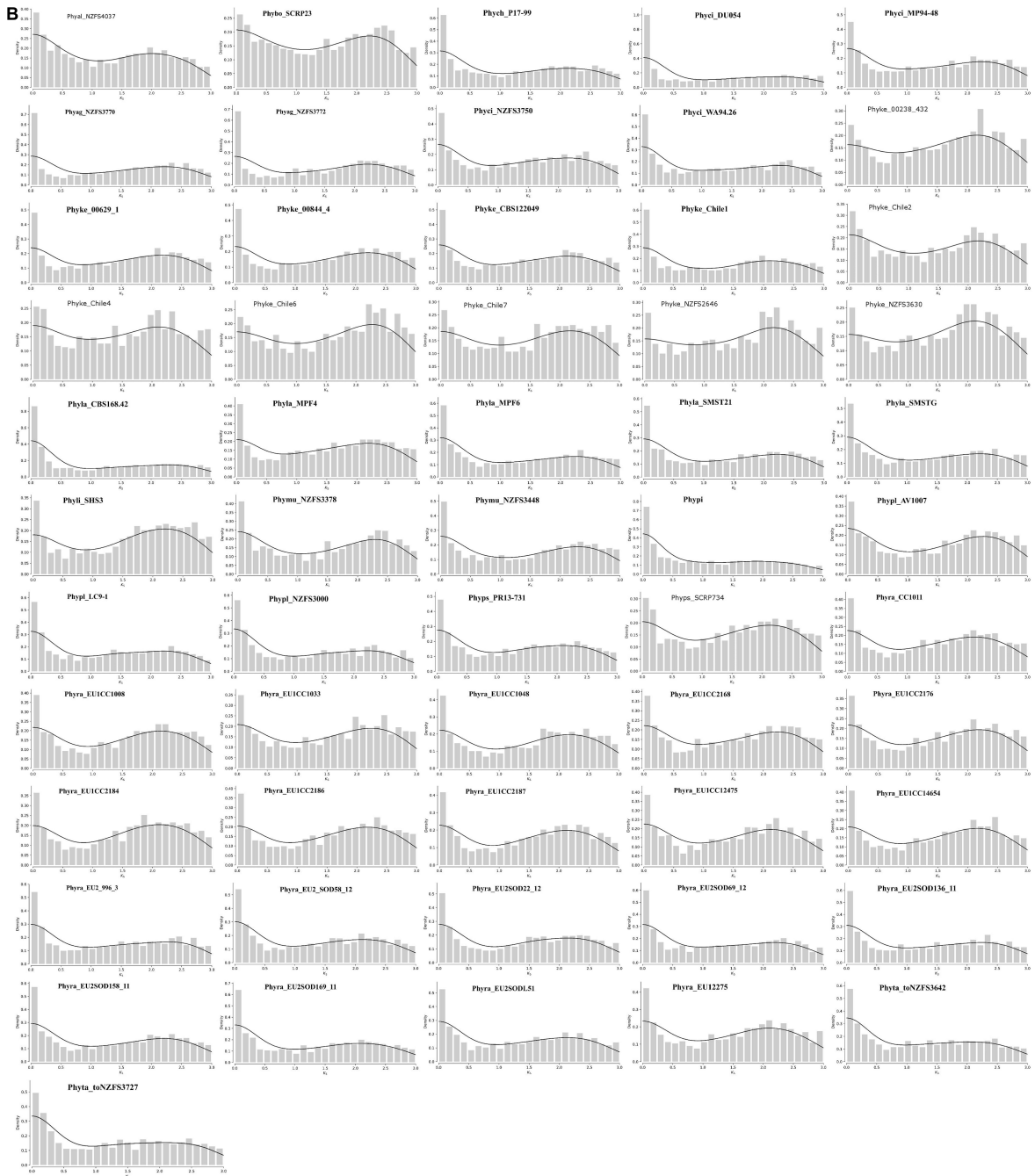


FIGURE 6 | K_S distribution of full paranomes. **(A)** Type-I whole genome duplication with KDEs of peaks in the K_S. Exponential decrease represents L-shaped curve that indicates whole genome duplication in ancestral time and loss of many duplicate genes. **(B)** Type-II whole genome duplication with KDE. Peaks in the K_S at 2–2.5 indicate increase of duplication frequency.

ramorum isolates from Europe. *P. cinnamomi* isolates GKB4 show type I WGD and have a genome size of 106 MB, whereas other isolates, i.e., DU054, MP94-48, NZFS3750, and WA94.26 have type II WGD and their genome size is nearly half the size of GKB4. *P. aleatoria* and *P. boehmeria* also have type II WGD, but two strong peaks were observed, which might be the due to the duplication happening in different time points.

WGDs can result due to autopolyploidy or by allopolyploidy, and both the events have been reported in *Phytophthora* (Redondo et al., 2015). WGD followed by gene loss plays a major evolutionary force to gene sub-functionalization and neo-functionalization in plants and animal systems (Huang et al., 2013). Our analysis suggests that the levels of genome duplication are largely

due to their genome localization in a specific geographical region and the selection pressure acting upon them. *Phytophthora* adapts to host-induced selection pressure by genome rearrangements and expansion mediated by repeats (Kamoun et al., 2015). Thus large-scale duplication events increase pathogen fitness in a given environment and specific host that is clearly evident from this analysis (Redondo et al., 2015).

CONCLUSION

The analysis of 128 *Phytophthora* genomes isolated from various geographical locations indicates that there is localized genome evolution and genome duplication. SSR motifs are preserved in an isolate-specific manner and can act as a unique identifier for a certain isolate. All the isolates of *Phytophthora* adhere to genome compartmentalization, where the core genes occur in compact regions of the genome. The infection-related genes and genes responsible for adaptive evolution, on the other hand, are localized in more repeat-rich regions amenable to rapid changes. All the annotated data and associated files are publicly deposited for community consumption. The browsable genomes and their annotations are available in www.eumicrobedb.org:3000.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

ST conceived the project. SD, KM, ST, AP, and AU carried out the data analysis. AU done the database design and upload with the help of AP. ST, KM, SD, and AU wrote the manuscript. All authors read and agreed on the contents of the manuscript.

FUNDING

This project was partially funded by CSIR-INDIA, MLP 134 to ST.

ACKNOWLEDGMENTS

KM, SD, and AU would like to thank University Grand Commission (UGC), Indian Council of Medical Research (ICMR), and Council of Scientific and Industrial Research (CSIR), respectively, for their fellowship.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.806398/full#supplementary-material

Supplementary Figure 1 | Overall pipeline used for secretome and effector prediction.

Supplementary Figure 2 | (A) Pearson's correlation coefficient of SSR number and Genome size. Result indicates number of SSRs positively correlated with genome size. **(B)** There was weak negative correlation between SSR density and genome size. **(C)** Number of SSR motifs shows weak negative correlation with genome size. **(D)** GC percentage of genome and SSR are positively correlated.

Supplementary Figure 3 | Heatmap depicts the percentage frequency of trinucleotide SSR group motifs in the coding sequences (CDS). There is a clear abundance of CAG/CTG, AAG/CTT motifs.

Supplementary Figure 4 | Heatmap depicts the in-frame cumulative frequency of trinucleotide SSR motifs in the coding sequences (CDS). There is a clear dominance of CAG, CTG, AAG motifs.

Supplementary Figure 5 | Effector prediction pipeline containing the number of proteins filtered out in each step. The first bar chart (bottom most) (light green) represents the genome sizes of 128 Phytophthora isolates studied. The second bar chart (ash) represents the total number of Open reading frames (ORFs) predicted by getorf function of EMBOSS package. These ORFs are further translated in one frame and the translated sequences were used for secretome prediction. The third bar chart (pink) shows the number of secretory proteins containing signal peptide (SP), based on SignalP v. 5.0b prediction. The fourth one (brown) shows the number of secretory proteins retained after the TMHMM analysis that does not contain any TransMembrane Helices (TMHs). The fifth one (violet) shows the number of secretory proteins that passed TargetP analysis. These retained proteins after the TargetP are used for effector prediction using EffectorO. The sixth bar chart (red) shows the number of predicted effectors for each genome. The seventh bar plot (green) represents the number of predicted RxLR effectors among the predicted effectors, predicted by homology searching. The eight (orange) and the last one (blue) show the number of effectors containing the CRN motif and the number of effectors containing the WYL domain respectively.

Supplementary Figure 6 | Box plots of FIRs of all genes, BUSCO genes and RxLRs in all the studied species. Here the positive FIR values were plotted with whisker = 0.2 parameter.

REFERENCES

Abril, J. F., and Castellano Hereza, S. (2019). *Genome Annotation*. Available online at: https://discovery.ucl.ac.uk/id/eprint/10058908/ (accessed August 2021).

Armenteros, J. J. A., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z

Basenko, E. Y., Pulman, J. A., Shanmugasundram, A., Harb, O. S., Crouch, K., Starns, D., et al. (2018). FungiDB: an integrated bioinformatic resource for fungi and *oomycetes*. *J Fungi* 4:39. doi: 10.3390/jof4010039

Baxter, L., Tripathy, S., Ishaque, N., Boot, N., Cabral, A., Kemen, E., et al. (2010). Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* 330, 1549–1551. doi: 10.1126/science.1195203

Biasi, A., Martin, F. N., Cacciola, S. O., di San Lio, G. M., Grünwald, N. J., and Schena, L. (2016). Genetic analysis of phytophthora nicotianae populations from different hosts using microsatellite markers. *Phytopathology* 106, 1006–1014. doi: 10.1094/PHYTO-11-15-0299-R

Birch, P. R. J., Armstrong, M., Bos, J., Boevink, P., Gilroy, E. M., Taylor, R. M., et al. (2009). Towards understanding the virulence functions of RXLR effectors of the oomycete plant pathogen *Phytophthora* infestans. *J. Exp. Bot.* 60, 1133–1140. doi: 10.1093/jxb/ern353

Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M., and Borodovsky, M. (2021). BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genom. Bioinform.* 3:lqaa108. doi: 10.1093/nargab/lqaa108

Buels, R., Yao, E., Diesh, C. M., Hayes, R. D., Munoz-Torres, M., Helt, G., et al. (2016). JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol.* 17:66. doi: 10.1186/s13059-016-0924-1

Byrt, P. N., Irving, H. R., and Grant, B. R. (1982). The effect of organic compounds on the encystment, viability and germination of zoospores of *Phytophthora cinnamomi*. *Microbiology* 128, 2343–2351. doi: 10.1099/00221287-128-10-2343

Cantarel, B. L., Korf, I., Robb, S. M. C., Parra, G., Ross, E., Moore, B., et al. (2008). MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 18, 188–196. doi: 10.1101/gr.6743907

Chepsergon, J., Motaung, T. E., and Moleleki, L. N. (2021). Core RxLR effectors in *phytopathogenic oomycetes*: a promising way to breeding for durable resistance in plants? *Virulence* 12, 1921–1935. doi: 10.1080/21505594.2021.1948277

Clark, T., Jurek, J., Kettler, G., and Preuss, D. (2005). A structured interface to the object-oriented genomics unified schema for XML-formatted data. *Appl. Bioinformatics* 4, 13–24. doi: 10.2165/00822942-200504010-00002

Cui, C., Herlihy, J. H., Bombarely, A., McDowell, J. M., and Haak, D. C. (2019). Draft Assembly of *Phytophthora capsici* from long-read sequencing uncovers complexity. *Mol. Plant. Microb. Interact.* 32, 1559–1563. doi: 10.1094/MPMI-04-19-0103-TA

da Fonseca, R. R., Albrechtsen, A., Themudo, G. E., Ramos-Madrigal, J., Sibbesen, J. A., Maretty, L., et al. (2016). Next-generation biology: sequencing and data analysis approaches for non-model organisms. *Mar. Genomics* 30, 3–13. doi: 10.1016/j.margen.2016.04.012

Derelle, R., López-García, P., Timpano, H., and Moreira, D. (2016). A phylogenomic framework to study the diversity and evolution of *stramenopiles (= heterokonts)*. *Mol. Biol. Evol.* 33, 2890–2898. doi: 10.1093/molbev/msw168

Dodds, P. N., and Rathjen, J. P. (2010). Plant immunity: towards an integrated view of plant–pathogen interactions. *Nat. Rev. Genet.* 11, 539–548. doi: 10.1038/nrg2812

Dong, S., Raffaele, S., and Kamoun, S. (2015). The two-speed genomes of filamentous pathogens: waltz with plants. *Curr. Opin. Genet. Dev.* 35, 57–65. doi: 10.1016/j.gde.2015.09.001

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.* 5, 435–445. doi: 10.1038/nrg1348

Engelbrecht, J., Duong, T. A., and Berg, N. V. D. (2017). New microsatellite markers for population studies of *Phytophthora cinnamomi*, an important global pathogen. *Sci. Rep.* 7:17631. doi: 10.1038/s41598-017-17799-9

Engelbrecht, J., Duong, T. A., Prabhu, S. A., Seedat, M., and van den Berg, N. (2021). Genome of the destructive *oomycete Phytophthora cinnamomi* provides insights into its pathogenicity and adaptive potential. *BMC Genomics* 22:302. doi: 10.1186/s12864-021-07552-y

Franceschetti, M., Maqbool, A., Jiménez-Dalmaroni, M. J., Pennington, H. G., Kamoun, S., and Banfield, M. J. (2017). Effectors of filamentous plant pathogens: commonalities amid diversity. *Microbiol. Mol. Biol. Rev.* 81:16. doi: 10.1128/MMBR.00066-16

Fulnečková, J., Ševčíková, T., Fajkus, J., Lukešová, A., Lukeš, M., Vlček, Č, et al. (2013). A broad phylogenetic survey unveils the diversity and evolution of telomeres in Eukaryotes. *Genome Biol. Evol* 5, 468–483. doi: 10.1093/gbe/evt019

Gao, R.-F., Wang, J.-Y., Liu, K.-W., Yoshida, K., Hsiao, Y.-Y., Shi, Y.-X., et al. (2021). Comparative analysis of *Phytophthora* genomes reveals oomycete pathogenesis in crops. *Heliyon* 7:e06317. doi: 10.1016/j.heliyon.2021.e06317

Gemayel, R., Cho, J., Boeynaems, S., and Verstrepen, K. J. (2012). Beyond junk-variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. *Genes* 3, 461–480. doi: 10.3390/genes3030461

Gonthier, P., Sillo, F., Lagostina, E., Roccotelli, A., Cacciola, O. S., Stenlid, J., et al. (2015). Selection processes in simple sequence repeats suggest a correlation with their genomic location: insights from a fungal model system. *BMC Genomics* 16:1107. doi: 10.1186/s12864-015-2274-x

Guo, W.-J., Ling, J., and Li, P. (2009). Consensus features of microsatellite distribution: microsatellite contents are universally correlated with recombination rates and are preferentially depressed by centromeres in multicellular eukaryotic genomes. *Genomics* 93, 323–331. doi: 10.1016/j.ygeno.2008.12.009

Guo, Y., Sakalidis, M. L., Torres-Londoño, G. A., and Hausbeck, M. (2021). Population structure of a worldwide *Phytophthora palmivora* collection suggests lack of host specificity and reduced genetic diversity in South American and Caribbean. *Plant Dis.* 105, 4031–4041. doi: 10.1094/PDIS-05-20-1055-RE

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., et al. (2009). Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 461, 393–398. doi: 10.1038/nature08358

Hannat, S., Pontarotti, P., Colson, P., Kuhn, M.-L., Galiana, E., La Scola, B., et al. (2021). Diverse trajectories drive the expression of a giant virus in the oomycete plant pathogen *Phytophthora parasitica*. *Front. Microbiol.* 12:662762. doi: 10.3389/fmicb.2021.662762

Hieno, M. A., Wibowo, A., Subandiyah, S., Shimizu, M., Suga, H., et al. (2019). Genetic diversity of *Phytophthora palmivora* isolates from Indonesia and Japan using rep-PCR and microsatellite markers. *J. Gen. Plant Pathol.* 85, 367–381. doi: 10.1007/s10327-019-00853-x

Huang, S., Ding, J., Deng, D., Tang, W., Sun, H., Liu, D., et al. (2013). Draft genome of the kiwifruit Actinidia chinensis. *Nat. Commun.* 4:2640. doi: 10.1038/ncomms3640

Ihaka, R., and Gentleman, R. (1996). R: a language for data analysis and graphics. *J. Comput. Graph. Stat.* 5:299. doi: 10.2307/1390807

Jiang, H., Hwang, H. W., Ge, T., Cole, B., Perkins, B., and Hao, J. (2019). Leucine regulates zoosporic germination and infection by *Phytophthora erythroseptica*. *Front. Microbiol.* 10:131. doi: 10.3389/fmicb.2019.00131

Jiang, R. H. Y., Tripathy, S., Govers, F., and Tyler, B. M. (2008). RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4874–4879. doi: 10.1073/pnas.0709303105

Kamoun, S., Furzer, O., Jones, J. D. G., Judelson, H. S., Ali, G. S., Dalio, R. J. D., et al. (2015). The Top 10 oomycete pathogens in molecular plant pathology. *Mol. Plant Pathol.* 16, 413–434. doi: 10.1111/mpp.12190

Karaoglu, H., Lee, C. M. Y., and Meyer, W. (2005). Survey of simple sequence repeats in completed fungal genomes. *Mol. Biol. Evol.* 22, 639–649. doi: 10.1093/molbev/msi057

Kassambara, A., and Kassambara, M. A. (2020). *Package "Ggpubr."* R Package Version 0. 1 6. Available online at: https://cran.microsoft.com/snapshot/2017-02-26/web/packages/ggpubr/ggpubr.pdf (accessed August 2021).

Katz, L., Griswold, T., Morrison, S., Caravas, J., Zhang, S., Bakker, H., et al. (2019). Mashtree: a rapid comparison of whole genome sequence files. *J. Open Sour. Softw.* 4:1762. doi: 10.21105/joss.01762

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315

Lagercrantz, U., Ellegren, H., and Andersson, L. (1993). The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Res.* 21, 1111–1115. doi: 10.1093/nar/21.5.1111

Lechner, M., Findeiss, S., Steiner, L., Marz, M., Stadler, P. F., and Prohaska, S. J. (2011). Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics* 12:124. doi: 10.1186/1471-2105-12-124

Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi: 10.1093/nar/gkab301

Mahfooz, S., Singh, S. P., Mishra, N., and Mishra, A. (2017). A comparison of microsatellites in *phytopathogenic aspergillus* species in order to develop markers for the assessment of genetic diversity among its isolates. *Front. Microbiol.* 8:1774. doi: 10.3389/fmicb.2017.01774

Mahfooz, S., Singh, S. P., Rakh, R., Bhattacharya, A., Mishra, N., Singh, P. C., et al. (2016). A comprehensive characterization of simple sequence repeats in the sequenced trichoderma genomes provides valuable resources for marker development. *Front. Microbiol.* 7:575. doi: 10.3389/fmicb.2016.00575

Mahfooz, S., Srivastava, A., Srivastava, A. K., and Arora, D. K. (2015). A comparative analysis of distribution and conservation of microsatellites in the transcripts of sequenced *Fusarium* species and development of genic-SSR

markers for polymorphism analysis. *FEMS Microbiol. Lett.* 362:fnv131. doi: 10.1093/femsle/fnv131

Malar, C. M., Yuzon, J. D., Das, S., Das, A., Panda, A., Ghosh, S., et al. (2019). Haplotype-Phased genome assembly of virulent phytophthora ramorum isolate nd886 facilitated by long-read sequencing reveals effector polymorphisms and copy number variation. *Mol. Plant. Microb. Interact.* 32, 1047–1060. doi: 10.1094/MPMI-08-18-0222-R

Marano, A. V., Jesus, A. L., de Souza, J. I., Jerônimo, G. H., Gonçalves, D. R., Boro, M. C., et al. (2016). Ecological roles of *saprotrophic Peronosporales (Oomycetes. Straminipila)* in natural environments. *Fungal Ecol.* 19, 77–88. doi: 10.1016/j.funeco.2015.06.003

Martens, C., and Van de Peer, Y. (2010). The hidden duplication past of the plant pathogen *Phytophthora* and its consequences for infection. *BMC Genomics* 11:353. doi: 10.1186/1471-2164-11-353

Mascheretti, S., Croucher, P. J. P., Vettraino, A., Prospero, S., and Garbelotto, M. (2008). Reconstruction of the sudden oak death epidemic in California through microsatellite analysis of the pathogen *Phytophthora ramorum*. *Mol. Ecol.* 17, 2755–2768. doi: 10.1111/j.1365-294X.2008.03773.x

McGowan, J., and Fitzpatrick, D. A. (2017). Genomic, network, and phylogenetic analysis of the oomycete effector arsenal. *mSphere* 2, e00408–e00417. doi: 10.1128/mSphere.00408-17

McGowan, J., and Fitzpatrick, D. A. (2020). Recent advances in *oomycete* genomics. *Adv. Genet.* 105, 175–228. doi: 10.1016/bs.adgen.2020.03.001

Morales-Cruz, A., Ali, S. S., Minio, A., Figueroa-Balderas, R., García, J. F., Kasuga, T., et al. (2020). Independent whole-genome duplications define the architecture of the genomes of the devastating West African cacao black pod pathogen *Phytophthora megakarya* and its close relative *Phytophthora palmivora*. *G3* 10, 2241–2255. doi: 10.1534/g3.120.401014

Morgante, M., Hanafey, M., and Powell, W. (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat. Genet.* 30, 194–200. doi: 10.1038/ng822

Nur, M., Wood, K., and Michelmore, R. (2021). EffectorO: motif-independent prediction of effectors in oomycete genomes using machine learning and lineage specificity. *bioRxiv* [Preprint]. Available online at: https://www.biorxiv.org/content/10.1101/2021.03.19.436227v1.abstract (accessed August 2021).

Olango, T. M., Tesfaye, B., Pagnotta, M. A., Pè, M. E., and Catellani, M. (2015). Development of SSR markers and genetic diversity analysis in enset (Ensete ventricosum (Welw.) Cheesman), an orphan food security crop from Southern Ethiopia. *BMC Genet.* 16:98. doi: 10.1186/s12863-015-0250-8

Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., et al. (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* 17:132. doi: 10.1186/s13059-016-0997-x

Panda, A., Sen, D., Ghosh, A., Gupta, A. C. M. M., and Prakash Mishra, G. (2018). EumicrobeDBLite: a lightweight genomic resource and analytic platform for draft *oomycete* genomes. *Mol. Plant Pathol.* 19, 227–237. doi: 10.1111/mpp.12505

Parada-Rojas, C. H., and Quesada-Ocampo, L. M. (2018). Analysis of microsatellites from transcriptome sequences of *Phytophthora capsici* and applications for population studies. *Sci. Rep.* 8:5194. doi: 10.1038/s41598-018-23438-8

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033

Raffaele, S., and Kamoun, S. (2012). Genome evolution in filamentous plant pathogens: why bigger can be better. *Nat. Rev. Microbiol.* 10, 417–430. doi: 10.1038/nrmicro2790

Redondo, M. A., Boberg, J., Olsson, C. H. B., and Oliva, J. (2015). Winter conditions correlate with phytophthora alni subspecies distribution in Southern Sweden. *Phytopathology* 105, 1191–1197. doi: 10.1094/PHYTO-01-15-0020-R

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the european molecular biology open software suite. *Trends Genet.* 16, 276–277. doi: 10.1016/S0168-9525(00)02024-2

Salzberg, S. L. (2019). Next-generation genome annotation: we still struggle to get it right. *Genome Biol.* 20:92. doi: 10.1186/s13059-019-1715-2

Schena, L., Cardle, L., and Cooke, D. E. L. (2008). Use of genome sequence data in the design and testing of SSR markers for *Phytophthora* species. *BMC Genomics* 9:620. doi: 10.1186/1471-2164-9-620

Schornack, S., van Damme, M., Bozkurt, T. O., Cano, L. M., Smoker, M., Thines, M., et al. (2010). Ancient class of translocated *oomycete* effectors targets the host nucleus. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17421–17426. doi: 10.1073/pnas.1008491107

Selkoe, K. A., and Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* 9, 615–629. doi: 10.1111/j.1461-0248.2006.00889.x

Seppey, M., Manni, M., and Zdobnov, E. M. (2019). BUSCO: assessing genome assembly and annotation completeness. *Methods Mol. Biol.* 1962, 227–245. doi: 10.1007/978-1-4939-9173-0_14

Srivastava, S., Avvaru, A. K., Sowpati, D. T., and Mishra, R. K. (2019). Patterns of microsatellite distribution across eukaryotic genomes. *BMC Genomics* 20:153. doi: 10.1186/s12864-019-5516-5

Stam, R., Jupe, J., Howden, A. J. M., Morris, J. A., Boevink, P. C., Hedley, P. E., et al. (2013). Identification and characterisation CRN effectors in phytophthora capsici shows modularity and functional diversity. *PLoS One* 8:e59517. doi: 10.1371/journal.pone.0059517

Stein, L. (2001). Genome annotation: from sequence to biology. *Nat. Rev. Genet.* 2, 493–503. doi: 10.1038/35080529

Stewart, S., Robertson, A. E., Wickramasinghe, D., Draper, M. A., Michel, A., and Dorrance, A. E. (2016). Population structure among and within iowa. missouri, ohio, and south dakota populations of *Phytophthora sojae*. *Plant Dis.* 100, 367–379. doi: 10.1094/PDIS-04-15-0437-RE

Studholme, D. J., McDougal, R. L., Sambles, C., Hansen, E., Hardy, G., Grant, M., et al. (2016). Genome sequences of six *Phytophthora* species associated with forests in New Zealand. *Genom Data* 7, 54–56. doi: 10.1016/j.gdata.2015.11.015

Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S., and McCouch, S. (2001). Computational and experimental analysis of microsatellites in rice *(Oryza sativa L.):* frequency, length variation, transposon associations, and genetic marker potential. *Genome Res.* 11, 1441–1452. doi: 10.1101/gr.184001

Tóth, G., Gáspári, Z., and Jurka, J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.* 10, 967–981. doi: 10.1101/gr.10.7.967

Tripathy, S., Deo, T., and Tyler, B. M. (2012). Oomycete transcriptomics database: a resource for *oomycete* transcriptomes. *BMC Genomics* 13:303. doi: 10.1186/1471-2164-13-303

Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R. H. Y., Aerts, A., et al. (2006). Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313, 1261–1266. doi: 10.1126/science.1128796

Vetukuri, R. R., Tripathy, S., Malar, C. M., Panda, A., Kushwaha, S. K., Chawade, A., et al. (2018). Draft genome sequence for the tree pathogen *Phytophthora plurivora*. *Genome Biol. Evol.* 10, 2432–2442. doi: 10.1093/gbe/evy162

Wang, X., and Wang, L. (2016). GMATA: an integrated software package for genome-scale SSR mining. marker development and viewing. *Front. Plant Sci.* 7:1350. doi: 10.3389/fpls.2016.01350

Wang, X., Yang, S., Chen, Y., Zhang, S., Zhao, Q., Li, M., et al. (2018). Comparative genome-wide characterization leading to simple sequence repeat marker development for Nicotiana. *BMC Genomics* 19:500. doi: 10.1186/s12864-018-4878-4

Wawra, S., Trusch, F., Matena, A., Apostolakis, K., Linne, U., Zhukov, I., et al. (2017). The RxLR motif of the host targeting effector AVR3a of *Phytophthora infestans* is cleaved before secretion. *Plant Cell* 29, 1184–1195. doi: 10.1105/tpc.16.00552

Whisson, S. C., Boevink, P. C., Moleleki, L., Avrova, A. O., Morales, J. G., Gilroy, E. M., et al. (2007). A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* 450, 115–118. doi: 10.1038/nature06203

Wickham, H. (2011). Ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* 3, 180–185. doi: 10.1002/wics.147

Yang, M., Duan, S., Mei, X., Huang, H., Chen, W., Liu, Y., et al. (2018). The *Phytophthora cactorum* genome provides insights into the adaptation to host defense compounds and fungicides. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-018-24939-2

Yang, X., Tyler, B. M., and Hong, C. (2017). An expanded phylogeny for the genus *Phytophthora*. *IMA Fungus* 8, 355–384. doi: 10.5598/imafungus.2017.08.02.09

Zhang, Q., Feng, R., Zheng, Q., Li, J., Liu, Z., Zhao, D., et al. (2019). Population genetic analysis of *phytophthora* parasitica from tobacco in Chongqing. *Southwestern China. Plant Dis.* 103, 2599–2605. doi: 10.1094/PDIS-05-18-0879-RE

Zwaenepoel, A., and Van de Peer, Y. (2018). Wgd-simple command line tools for the analysis of ancient whole-genome duplications. *Bioinformatics* 35, 2153–2155. doi: 10.1093/bioinformatics/bty915

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership